



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

문학석사 학위논문

# Source of Moral Motivation

## 도덕적 동기의 원천

2016년 2월

서울대학교 대학원

철학과 서양철학 전공

김 대 현

# Source of Moral Motivation

지도교수 김 현 섭

이 논문을 문학석사 학위논문으로 제출함

2016년 2월

서울대학교 대학원  
철학과 서양철학전공  
김 대 현

김대현의 문학석사 학위논문을 인준함  
2016년 2월

위 원 장 \_\_\_\_\_ (인)

부위원장 \_\_\_\_\_ (인)

위 원 \_\_\_\_\_ (인)

## Abstract

# Source of Moral Motivation

Daehyun Kim

Department of Western Philosophy

The Graduate School

Seoul National University

The purpose of this thesis is twofold. One, I look to provide an unconventional yet plausible externalist account of the relationship between moral judgment and motivation. Two, in defending my externalist account, I present arguments against its counterpart internalist view. Regarding one, the framework of my position rejects both internalism and the Humean theory of motivation. In other words, I propose that the connection between moral judgment and motivation is *intrinsic* and *contingent* at the same time – that is, while the source of moral motivation is internal to moral judgment, it is not necessarily the case that moral judgments will always motivate. Regarding two, I specifically target an internalist account which I believe adequately contrasts to my position: Michael Smith’s cognitivist internalist view.

Smith argues that a fully rational agent would desire accordingly to the normative reasons she accepts. Therefore, internalism is true and any agent who acknowledges some normative reason

for action yet is not motivated appropriately is in some sense practically irrational. I question the idea that moral judgments are constrained by norms of rationality by providing arguments aimed to show that moral judgments are not subject to requirements of consistency, a key tenet of rational behavior. I further argue that internalist views such as Smith's make the mistake of conflating the action-guiding nature of morality to motivation; that is, what it means for something to be normative and for that something to be motivating.

In asserting my externalist position, I argue that within the deliberative processes of making a moral judgment, there are at least two distinct forms of commitment made understood as judging what is (morally) best and judging to do that which is (morally) best. This distinction between two ways of holding a moral judgment is then elaborated upon to show that it is the latter which is most aptly connected to motivation. Finally, I argue that the operative state within such moral judgments responsible for motivation is a higher degree of commitment by the agent whereby she turns an essential *impersonal moral imperative* into a *personal imperative*.

I conclude by addressing an internalist argument against the plausibility of an amoralist known as the inverted commas objection. Smith's version of the objection claims that an amoralist lacks 'sufficient mastery of moral concepts'. I respond by arguing that Smith's account fails to accommodate a certain type of amoralist: one who has recently become that way. Additionally, I supplement my argument by claiming that the inverted commas objection conflicts with the part of our intuition that is influenced primarily by observational evidence.

Keywords: Cognitivism, Externalism, Humean Theory of Motivation, Internalism, Inverted Commas Objection, Moral Judgment, Motivation, Non-Cognitivism

Student Number: 2012-23886

# Table of Contents

Abstract (English) / i

Introduction / 1

I. An Initial Acquaintance with the Topic / 8

1.1. Taxonomy of Motivational Internalism / 8

1.2. Three Divergent Views on Motivational Judgment Internalism/Externalism / 10

1.3. Regarding the Efficacy of Motivation Following Moral Judgment / 12

1.4. Motivational Judgment Externalism / 14

II. A Defense Against the Weak Internalist View / 16

2.1. Smith's Cognitivist Internalist View / 16

2.2. On the Relationship Between Rationality and Motivation / 22

2.3. The Externalist Claim / 27

2.4. Action-Guiding Not the Same as Motivating / 32

III. The Source of Moral Motivation / 36

3.1. Non-Humean Externalism / 36

3.2. A Starting Distinction Between Two Types of Moral Judgments / 43

3.3. Moore's Paradox as an Analytical Tool / 46

3.4. The Operative State Within Moral Judgments Responsible for Motivation / 48

IV. Can Amoralists Really Make Genuine Moral Judgments? / 53

4.1. The Inverted Commas Objection / 54

4.2. Insincere Moral Statements / 56

4.3. Non-Cognitivist Versus Cognitivist's Use of the Inverted Commas Objection / 58

4.4. Smith's Inverted Commas Objection /	61
4.5. A Rebuttal to Smith's View /	64
4.6. An Amoralist's Judgment is the Same in <i>Kind</i> as Those of a Moral Agent's /	67
Closing Remarks /	69
References /	74
Abstract (Korean) /	77

## **Introduction**

Moral motivation is generally understood as a type of normative motivation – a force that seems to compel us towards wanting to act in accordance with our moral judgment. Often times this phenomenon is explained by the idea that when we judge something to be morally right or good, not only does the moral judgment then give us a reason to act accordingly or such action seems rational to take, but we also therefore tend to be moved. One of the more interesting aspects of moral motivation is that it can often conflict with an agent's non-moral values or other preferences in that, frequently, there seems to be an opposition between self-interested desires and the norms of morality (or what one thinks she ought to do). In fact, it often seems to be the case that one has to make sacrifices of one's desires in order to follow the moral good. Furthermore, what seems to complicate our goal of trying to understand moral motivation is that sometimes there seems to be such a reliable connection between moral judgments and motivation; while at other times, there seems to be such spectacular failure of motivation following moral judgment. Trying to explain this inconsistency is a crucial aspect of motivational research and has led to the two opposing views of motivational internalism and externalism.

One may pose the question why we should care about the relationship between moral judgment and motivation in the first place. Such a person might think 'so what?' if motivation necessarily arises whenever a moral judgment is made; or if motivation is contingent upon some other, external factor outside of moral judgment. Why should that affect my personal beliefs regarding what is right or wrong? To provide one answer to this sort of query, if nothing else, understanding the relationship between moral judgment and motivation has a direct impact in further understanding the entire meta-ethical landscape. And this is important because there is no doubt a great amount of disagreement in meta-ethics.



In the introductory chapter of his book, *The Moral Problem*, Michael Smith makes precisely this point in considering why there is so much disagreement in the domain of meta-ethics. He suggests that the reason for disagreement can be traced back to two distinctive features of morality which carry implications that pull in opposite directions; with the conflict being further exasperated by what he calls ‘the standard picture of human psychology – a picture we owe to Hume (1888)’.<sup>1</sup> Smith calls this ‘the moral problem’ and provides a succinct layout of the conflict by presenting each of the two features of morality he mentions plus the Humean account of human psychology in the form of three apparently inconsistent propositions regarding the relationship between moral judgment and motivation.

1. Moral judgments of the form ‘It is right that I  $\Phi$ ’ express a subject’s beliefs about an objective matter of fact, a fact about what is right for her to do.
2. If someone judges that it is right that she  $\Phi$ s then, *ceteris paribus*, she is motivated to  $\Phi$ .
3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume’s terms, distinct existences.<sup>2</sup>

The first assumption is simply the cognitivist view that judgments are beliefs, or an agent’s representation of certain aspects of the world. The second is a version of the motivational internalist’s claim that moral judgments necessarily entail motivation.<sup>3</sup> The third claim reflects the Humean theory of motivation. This view expresses the general Humean thought that non-cognitive

---

<sup>1</sup> Smith 1994: 7.

<sup>2</sup> Ibid: 12.

<sup>3</sup> To be more accurate, this is a particularly weak form of motivational internalism. To see this, consider the following two forms of weak internalism. The first asserts a defeasible connection between moral judgment and action: though moral judgments necessarily motivate, this moral motivation can be overridden by countervailing motivation and thus not lead to any action corresponding to the moral judgment. In contrast, the second (even weaker) form, which is reflective of Smith’s view, asserts a defeasible connection between moral judgment and motivation itself. Thus, Smith’s view can accommodate the idea that an agent making a moral judgment may not even have any motivation at all, which seems to make his view particularly amenable to externalism.

states, such as desires or pro-attitudes, are necessary psychological states for motivation and that cognitive mental states such as beliefs are conceptually or modally independent from non-cognitive mental states such as desires or pro-attitudes. Even though not everyone accepts each of these three claims, nonetheless, the claims have some intuitive appeal and show a certain degree of plausibility, at least individually. However, combined it is obvious that there is an inconsistency; that not all three together can be true.<sup>4</sup> Therefore, in order to avoid a contradiction, we must reject at least one of the three claims.

Non-cognitivists reject claim one, asserting that a moral judgment is some form of a non-cognitive attitude. The motivational externalist rejects claim two, saying that a moral judgment (represented as a cognitive state), at least in and of itself, cannot motivate. Rejecting assumption three is known as non-Humeanism.<sup>5,6</sup> Non-Humeanism can take two forms. Some who reject assumption three propose that beliefs in themselves can sufficiently motivate without requiring any non-cognitive, desire-like state. The alternate form concedes that both a belief-like state and a desire-like state are necessary for motivation, but that belief-like states and desire-like states have

---

<sup>4</sup> A simplified explanation of the inconsistency is as follows. (1) tells us that moral judgments express belief-like mental states. (2) claims for a necessary connection between such belief-like states and motivation. However, (3) claims that a desire-like state is necessary for motivation but also that there is no necessary connection between a belief-like state and a desire-like state, hence rebutting the conjunction of (1) and (2). Strictly speaking however, it could be argued that there is no contradiction in Smith's version of the three premises. This is because, as partly explained in the previous footnote, Smith's version of internalism (2) has a *ceteris paribus* clause, that belief and motivation can come apart. In order for there to be a contradiction, (2) would most likely have to be read as: (2\*) *Necessarily*, if someone judges that it is right that she  $\Phi$ s, then she is motivated to  $\Phi$ .

<sup>5</sup> Many authors use the term 'anti-Humeanism' instead of 'non-Humeanism' to reflect a contrasting view to the Humean theory of motivation. If there is a distinction between anti-Humeanism and non-Humeanism, I would imagine that the anti-Humean position is more directly opposed to what is entailed by the Humean theory of motivation, whereas the non-Humean position is not so much concerned with going against the Humean view, but rather offers a different view that is simply not in line with Humeanism. I am not quite sure if this is correct or if there is any more to it than that. But if we grant this distinction here, because I believe that the stronger position of anti-Humeanism might restrict what can be entailed or represented by the view I look to forward, I opt to use the non-Humean terminology to depict the position contrasting Humeanism in my thesis.

<sup>6</sup> David Brink notes that internalists who accept cognitivism (those who accept Smith's assumptions one and two) but deny the third assumption can also be called 'rationalists'.

some conceptual connection to each other in that a normative belief can produce the desire-like state necessary for motivation.

Now as I have said, in order to hold a coherent picture, it suffices to simply reject one of the three views. However, I reject both assumptions two and three (at least loosely stated). But a mere denial of internalism and Humeanism is not what I set out to do in this thesis. Rather, I look to ‘throw my hat’ into the internalist/externalist debate by offering an externalist account which tries to expand the limits to how we should understand the relationship between moral judgment and motivation unlike what has been traditionally suggested by either internalists or externalists so far.<sup>7</sup> Therefore, I simply take externalism and non-Humeanism as the framework from which I present my position: and that is to argue for an *intrinsic yet contingent connection* between moral judgment and motivation. In other words, I argue that while the source of motivation is connected to the moral judgment – as opposed to being external to it – it is a contingent relationship nonetheless.

The main focus of my thesis is to provide an externalist account of the relationship between moral judgment and motivation, one that stands against Smith’s internalist account. Although admittedly this might be too simplistic, it could be understood that the main difference between Smith’s account and a more generic internalist view is, one, regarding the nature of moral judgments – that they are cognitive mental states rather than non-cognitive – and, two, the role of practical rationality insofar as motivation is concerned. Thus, chapter II focuses on laying out precisely what Smith’s cognitivist internalist account amounts to, and then following with externalist arguments against Smith’s view. More specifically speaking, after providing an account of Smith’s view, I argue against the necessary connection between rationality and motivation, clarify what the externalist account is claiming, as well as provide an account of what I think is a

---

<sup>7</sup> Of course, I do not wish to imply that therefore my proposed view is a completely new and original position regarding this topic. Rather that this is not a commonly held position by most philosophers in the field.

mistake the internalists make in their assertion for the necessary connection between moral judgments and motivation. I claim that they conflate the ‘action-guiding’ nature of morality with motivation: that is, between what is meant for something to be normative and for that thing to be motivating.

I mentioned that I also reject the Humean theory of motivation. The non-Humean aspect of my thesis should be understood as a mere secondary consequence to my internalist/externalist discussion, reflected by my stance. I provide a cursory explanation in regard to this throughout chapter III. However, I do not go into much detail arguing for non-Humeanism as I think that would dilute from the main purpose of my paper, which is to discuss motivational internalism/externalism. The overall objective of chapter III will focus on developing my externalist position of the intrinsic yet contingent relationship between moral judgment and motivation.

The manner by which I look to support my externalist position is by first proposing that there is more than one way an agent can hold a moral belief (which is reflected by moral judgments of the forms ‘ $\Phi$  is right’ and ‘I/one ought to  $\Phi$ ’). The purpose of proposing two ways of holding a moral belief is to support the externalist intuition that there are some propositional forms of moral judgments that seem to lack motivational force, from an agent’s psychological perspective. Specifically speaking, although I acknowledge that moral judgments of the form ‘ $\Phi$  is right’ are logically connected to moral judgments of the form ‘I/one ought to  $\Phi$ ’, I argue that they differ in respect to their relationship to motivation. What this means is that, while ultimately the connection between either propositional form of a moral judgment to motivation can be understood as being a contingent one, I will argue that the morally motivated agent is one who perceives her moral judgment as not only being in the form of a moral imperative (so the ‘I/one ought to  $\Phi$ ’ form), but furthermore as a *personal imperative*. That is, I argue that motivation follows a moral judgment

when the agent *personally engages* with the normative content of her moral judgment, which entails a difference in the level of commitment by the agent towards the normative content.

The purpose of chapter IV is to discuss the ‘inverted commas objection’: a prominent internalist argument against the idea that an agent can make a sincere moral judgment without being motivated. Since my externalist position posits a contingent connection between moral judgment and motivation – and so allowing for the possibility that an agent can make a moral judgment without being motivated – this objection is one that holds features regarding the relationship between moral judgment and motivation that necessarily contradict my view. Therefore, by providing compelling arguments against the inverted commas objection, I will be able to address the criticism against the plausibility of someone who can make a moral judgment and not be motivated, and hence strengthening my own position.

The first half of chapter IV will consist in a discussion of the inverted commas argument in its generic form. I introduce what this argument amounts to in its most general sense and offer initial concerns over the formulation of the argument. I question whether the foundation of the objection, that the necessary connection between moral judgment and motivation is a conceptual truth, can really be asserted as a conceptual truth. I then present an example of what the inverted commas objection could amount to and how the externalist presentation of the plausibility of an ‘amoralist’ differs from such an objection.

In the latter half of chapter IV, I present a specifically cognitivist internalist version of the inverted commas objection, namely Smith’s. The reason for this is that since my externalist position can be understood as considering moral judgments to be belief-like cognitive states, I can respond to a non-cognitive internalist charge of inverted commas usage against me, not by defending against the inverted commas objection specifically, but by simply denying that moral judgments should be understood as non-cognitive (conative) states that are necessarily motivating.

But a cognitivist internalist's charge of making moral judgments in an inverted commas sense takes as its starting point the same position as mine regarding the nature of moral judgments. Therefore, such a cognitivist internalist's charge of inverted commas usage is a much more unambiguously applicable argument against my position; and it is such a position that I need to specifically address.

Therefore the latter half of chapter IV will mainly comprise of laying out what Smith's inverted commas objection entails, followed by my argument against his claims. In the last section of this chapter, I expand on a specific aspect of Smith's version of the inverted commas objection that seems to be particularly amicable towards my proposed account of being able to hold moral beliefs in more than one way. The aspect in question refers to the possibility that an amoralist's judgment is not different in *kind* from that of a moral agent's, but might instead just be missing some element that is otherwise present in a moral agent's judgment. I suggest that this missing element is precisely the agent engaging with the moral content, understood as a difference in the level of commitment by an agent towards the normative content of the moral judgment.

Before delving into these three chapters that comprise the main portion of my thesis, in the upcoming first chapter, I provide an initial acquaintance to the topic by going over variations of motivational internalism, as well as providing an explication of motivational judgment externalism. The purpose of chapter I is twofold. For one, I look to get readers who are not yet quite familiar with this subject matter a chance to gain some background knowledge before going into my positive account. The second purpose is to allow the more knowledgeable reader to transition more smoothly into my proposal. However, even if there is some confusion after reading this chapter – because there are so many distinctions and variety of sub-categories one has to be aware of – it should still be possible to follow along the more important chapters II through IV without further information.

## **I. An Initial Acquaintance with the Topic**

Trying to provide a comprehensive landscape of all the available varieties of views involved in the debate on the relationship between moral judgment and motivation is an improbable task for this thesis. Luckily, for the dual purpose of providing a general acquaintance to the topic, as well as for gaining a better understanding of my view, it suffices to sketch the rough outline that comprises this chapter. However, I note once again that the following outline should not be considered anywhere near all-inclusive for what is entailed by motivational internalism and externalism.

So to begin with, consider one common perception regarding moral judgments – that they seem to be of such a nature that when an agent utters a statement to the effect ‘it is right to  $\phi$ ’, she could not sincerely make such a statement without being motivated in some sense by that judgment herself. As mentioned in the introduction, this idea is called motivational internalism. Internalism simply means that there is some internal (necessary) connection between moral judgments and motivation. However, there are many different variations of motivational internalism – each with the aim of laying out a more conceptually plausible picture regarding moral motivation – that need to be distinguished. For example, is the internalist asserting that a moral judgment is, in and of itself, a motivational state; motivating due to psychological factors in an agent; or because moral judgments contain moral content which necessarily have motivating properties? These are just a few of the myriad internalist ways we can understand the relationship between moral judgments and motivation.<sup>8</sup>

### **1.1. Taxonomy of Motivational Internalism**

---

<sup>8</sup> It should be inferred that for any particular version of internalism there is a counterpart externalist view available.

The first type of variation distinguishes between different types of sources that are responsible for motivation when an agent makes a moral judgment (whatever the nature of a moral judgment is). The following various types of motivational internalist views are not necessarily considered to be at odds with each other but are rather just completely independent views. The views in question are called *existence internalism*, *internalism about reasons* (or *reasons internalism*), and *judgment internalism*.

The term existence internalism was developed by Stephen Darwall.<sup>9</sup> The underlying idea is that the normative status of a certain moral fact or entity – the contents of a moral judgment – necessarily motivates any agent that gains cognitive access to it. More specifically, this view holds that it is part of the makeup of the moral fact or object that it possesses the power or property to motivate. For one, this view clearly supports *moral realism*, that metaphysical entities known as moral facts or truths do exist in this world. For another, because this view holds that motivation necessarily arises regardless of any psychological impediments from the agent, it is a particularly extreme moral realist view. Existence internalism is often understood in terms of Plato's 'Form of the Good', and it is this version of internalism which J.L. Mackie refers to when rejecting moral realism in his well-known 'Argument from Queerness' objection.<sup>10</sup>

Reasons internalism is generally understood as a form of existence internalism and asserts that (i) a moral judgment gives an agent some reason in favor of it, and (ii) that the reason will necessarily give the agent some motivation to act. This view is credited to Bernard Williams who argues for a conceptual connection between moral action and what he calls a person's 'subjective motivational set'. According to Williams, a person's subjective motivational set – which is primarily discussed in terms of desires – includes many elements such as 'dispositions of evaluation,

---

<sup>9</sup> Darwall 1992: 155.

<sup>10</sup> Mackie 1997: 38-42.



patterns of emotional reaction, personal loyalties, and various projects'.<sup>11</sup> Furthermore, a person's subjective motivational set is not static; deliberative processes can have all sorts of effects on it in terms of addition, subtraction and revisions of elements contained in the set. The basic argument by Williams for reasons internalism is that when an agent believes herself to have a reason for action, she will necessarily be motivated because such reasons can only come about due to something in the agent's 'subjective motivational set' which already contains some desire regarding the content of the reason.

Judgment internalism, on the other hand, states that an agent cannot sincerely make a moral judgment without having some motivation corresponding to the judgment: that it is part of the psychology of an agent and the nature of moral judgments that when an agent makes a sincere moral judgment, she can only be doing so if she is also motivated accordingly. One way to differentiate between this view and existence internalism is that, whereas existence internalism is more of a metaphysical view regarding certain motivating properties that moral content possess, judgment internalism relies more on psychological explanations focusing on the concept of (moral) judgment itself and how to properly understand the idea of an agent making a judgment. It is this last view of *motivational judgment internalism* that is pertinent for this paper and will prominently be dealt with in the discussions to follow.

## 1.2. Three Divergent Views on Motivational *Judgment* Internalism/Externalism

While section 1.1 provided a distinction between different possible types of sources of moral motivation, the distinction in this section is specific to the third type – judgment internalism – and is between two different concepts on the *nature* of a moral judgment. The two views can be understood as *cognitivism* and *non-cognitivism* regarding the nature of moral judgments and they

---

<sup>11</sup> Williams 1981: 105.

present quite a divergent understanding of the relationship between moral judgment and motivation. Add to that the externalist picture, and we have three differing views in the debate on the relationship between moral judgment and motivation: *cognitive internalism* (CI), *non-cognitive internalism* (NCI), and *externalist realism* (ER).<sup>12</sup> Unlike in 1.1 where the three views were not at odds with each other and were rather three independent views, the three views mentioned here are directly opposed to each other in one way or another. Therefore the best way to understand each view (as well as the difference between cognitivism and non-cognitivism) is by doing a direct compare and contrast between the three. I start by comparing and contrasting CI and NCI.

As one might guess CI and NCI share in common the ‘internalist’ aspect of the relationship between moral judgment and motivation. What this means is that both views consider motivation to be intrinsically, or necessarily, connected to a moral judgment. Motivation will necessarily be entailed when an agent makes a sincere moral judgment. The way the two views differ is in the mechanism involved in how such motivation is necessarily connected to moral judgment. First, CI asserts that moral judgments are belief-like cognitive mental states. What this entails is that CI understands moral judgments as being fact-stating propositions, at least from the agent’s point of view. This does not mean that every moral judgment reflects a moral truth substantively, but rather that the content of the moral judgments are apt for evaluations of truth or falsity. Furthermore, CI understands such belief-like cognitive states to be able to motivate and do so necessarily. Thus, CI asserts that a belief-like state is sufficient somehow in guaranteeing moral motivation that corresponds to the moral belief.

Opposingly, NCI asserts that moral judgments represent non-cognitive (or conative) mental states such as desires or pro-attitudes. Since non-cognitive mental states are not fact-stating, the moral judgments expressed are not subject to truth or falsity. Also, due to the very nature of

---

<sup>12</sup> The term ‘externalist realism’ comes from Brink (1986).

conative states – they are considered to be motivationally efficacious – NCI is a view that serves as a natural explanation for why motivation necessarily has to follow a moral judgment. That is, if moral judgments simply are (either conceptually) or express (as representations of) non-cognitive mental states such as desires or pro-attitudes, then it just follows that they are motivational states.

In contrast to both CI and NCI, ER considers the relationship between moral judgment and motivation to be contingent. It is not necessarily the case that an agent making a moral judgment will be motivated accordingly. ER compares to CI in that both positions take the nature of moral judgments to be belief-like cognitive states. However, unlike CI, ER denies the possibility that a belief-like state can sufficiently motivate on its own. ER compares to NCI in that both positions advocate for some non-cognitive desire-like state as the crucial element necessary for motivation to occur. However, unlike NCI, which has the desire-like state built into the moral judgment, ER considers the desire-like state to be external to the moral judgment and not a necessary occurrence when a moral judgment is made. This then accounts for why NCI sees the relationship between moral judgment and motivation as a necessary one, whereas ER sees the relationship only as contingent.<sup>13</sup>

### 1.3. Regarding the Efficacy of Motivation Following Moral Judgment

So far I have given distinctions based on different possible types of sources of moral motivation and distinctions based on the nature of moral judgments themselves. In this section I lay out an important distinction between various degrees of strength, or to what extent, motivation is necessarily entailed from a moral judgment.

---

<sup>13</sup> Since my position supposes non-Humeanism, so arguing that moral judgments are belief-like cognitive states, I am explicitly narrowing the range of views to exclude any non-cognitivist position. Thus any arguments from and against internalism in this thesis should be understood as specifically addressing the cognitivist internalist position unless noted otherwise.

Motivational judgment internalism can be broadly divided into two categories of *strong* and *weak* motivational internalism. Within the strong class of motivational judgment internalism, there is a further distinction that can be made. The first strong view is simply the idea that moral judgments necessarily motivate *simpliciter*. That is, this view inevitably pushes the motivation entailment all the way to yielding action and is therefore a particularly extreme view indeed. We can call this view *Strong Act (Motivational Judgment) Internalism* (SAMJI) and is differentiated from *Strong Motivational (Judgment) Internalism* (SMJI), as phrased by Nick Zangwill.<sup>14</sup> SMJI contrasts with SAMJI in that it is a view that deals only with motivation rather than with both motivation and action. This view states that moral motivation overrides all other motivations. That is, even though an agent may fail to actually act upon the moral motivation – perhaps ‘due to failures of rationality or opportunity’ – nonetheless, the motivation that arises from a moral judgment is the agent’s strongest motivation. In contemporary philosophy, neither of these views that fall under the ‘strong’ category are widely accepted for: (i) it seems clear that at least sometimes, we do not act according to what we judge we morally ought to do; and (ii) for being too restrictive in that it is unable to account for cases such as weakness of will, akrasia, mental/psychological disorders or impairments, and so on.

Thus, generally proposed concepts of judgment internalism fall in the ‘weak’ category which we can call *Weak (Motivational Judgment) Internalism* (WMJI). WMJI is simply the view that while moral judgments necessarily entail corresponding motivation, this motivation may not be the strongest and can be overridden by countervailing motivations. ‘Countervailing motivations’ is to be understood broadly as not only simply being overwhelmed by conflicting desires, but also encompassing forms of practical irrationality such as suffering from some form of weakness of will, akrasia, mental/psychological disorders or impairments, and the like. What is important is that,

---

<sup>14</sup> Zangwill 2008: 93-94.

even though weak internalists thus admit that these countervailing motivations may lead to an agent not being moved to *act* upon their moral motivations, nonetheless such moral motivation does exist even to the slightest degree when an agent makes a moral judgement.

However, as I had briefly mentioned in footnote 3, Smith's version of weak motivational internalism is even weaker than this general weak internalist view. The main difference between Smith's weak internalist view and WMJI can be understood as that, for Smith, when moral motivation is overridden by said countervailing motivations, the connection between moral judgment and motivation can be severed completely. In other words, moral judgments motivate *ceteris paribus* – the relationship between moral judgments and motivation is a *defeasible* one. I will lay out in more detail what Smith's weaker internalist view entails in section 2.1.

#### 1.4. Motivational Judgment Externalism

The general externalist intuition is illuminated in the following phrase by Christine Korsgaard, “a conjunction of moral comprehension and total unmotivatedness is perfectly possible: knowledge is one thing and motivation another.”<sup>15</sup> In other words, the idea that forms of (moral) knowledge and motivation are necessarily connected is a view that needs to be argued for by the internalist and shown to be the case. The agent who reflects the position described by Korsgaard here goes by many names in the literature such as ‘the cynic’, ‘the skeptic’, ‘the indifferent or listless person’, or ‘the amoralist’. While there are arguably conceptual differences on a more acute level, for the most part I will refer to *the amoralist* as encompassing all the previously mentioned titles.

On the most basic level, the core of any externalist view maintains a rejection of the necessity condition insofar as the connection between moral judgment (or the moral content/reasons)

---

<sup>15</sup> Korsgaard 1986: 156. Korsgaard herself prescribes to internalism and, in this passage, is simply offering up a description of externalism.

and motivation is concerned. Any connection between moral judgment and motivation is asserted as only being a contingent one. How this contingent connection is elaborated on can take various forms. Perhaps the most common externalist method is one which posits that some form of a conative state (such as a desire), which is external to the moral judgment, is necessary in order for motivation to occur. This is often referred to as the agent's 'desire to be moral'.<sup>16</sup> This type of argument emphasizes that the source of moral motivation does not reside in the moral judgment itself but outside of it, as reflected by Thomas Nagel's statement that "externalism is compatible with the view that... motivation is always present – so long as its presence is not guaranteed by moral judgments themselves, but by something external to ethics."<sup>17</sup>

Also, the externalist position does not need to argue for the possibility that an amoralist is never motivated by the moral judgments she makes. The externalist needs only to argue that, at least on a single occasion, it is conceptually possible for an amoralist to make a sincere moral judgment without being correspondingly motivated. What this also entails is that the internalist position does not become vindicated simply by providing a successful argument that amoralists do not in fact exist. Rather, the internalist must show that it is *conceptually impossible* for an amoralist to exist. However, just as the internalist is faced with the burden of giving an argument for the conceptual impossibility of the existence of an amoralist, the externalist has to provide a compelling argument for the conceptual possibility of an amoralist – someone who can sincerely make a moral judgment without having any sort of corresponding motivation. Many externalists begin by noting the following. What is imaginable is conceptually possible. We can imagine an amoralist, therefore the amoralist must be conceptually possible. Thus, internalism must be false and externalism must be true.<sup>18</sup> This is not a good argument as every step of the reasoning process can

---

<sup>16</sup> Cf. Brink 1989, Railton 1986, Svavarsdóttir 2006.

<sup>17</sup> Nagel 1970: 7.

<sup>18</sup> Shafer-Landau 2003: 146.

be questioned. Internalists right away reject this argument since it goes against what they consider as a conceptual truth regarding moral judgments; that they necessarily motivate. Furthermore, this does nothing to illuminate whether such a supposed amoralist is actually making a *sincere* moral judgment or not. Internalists explain such an amoralist away by arguing that they do not make sincere moral judgments, but are rather simply using moral judgments in an ‘inverted commas’ sense.<sup>19</sup> I come back to the inverted commas objection in chapter IV as I believe that addressing this objection plays a vital role in defending my view. Either way, the use of the putative amoralist has aided externalists in their attempts to unseat internalism as an *a priori* constraint on meta-ethical theorizing.

## II. A Defense Against the Weak Internalist View

### 2.1. Smith’s Cognitivist Internalist View

I started this thesis by citing Smith’s layout of ‘the moral problem’. Smith’s personal meta-ethical position can be summarized as being a combination of precisely the three views which he claimed cannot be held together without contradiction: (1) moral cognitivism, (2) motivational internalism, and (3) the Humean theory of motivation. Nevertheless, Smith develops a sophisticated and comprehensive account throughout his book in order to resolve the inconsistency without having to reject any of the three positions and it is fair to say that his account has a great amount of appeal. However, I certainly cannot recapitulate his overall account here as I do not have the space for it and it would not do justice to his view anyhow. Instead, I focus on just the

---

<sup>19</sup> *Inverted commas* refers to R.M. Hare’s idea that when an amoralist utters a moral judgment such as ‘ $\phi$ -ing is right’, she is not really making a moral judgment, but rather simply saying that ‘ $\phi$ -ing is *right*’ where ‘right’ is understood as what the amoralist believes society, or other people, conceive of as right and not based on the agent’s own sincere understanding of the term. See Hare 1952: 124-126.

internalist component of his overall view. I start out by presenting Smith's internalist view in its most basic form, the way it is offered by Smith himself:

If an agent judges that it is right for her to  $\Phi$  in circumstances C, then either she is motivated to  $\Phi$  in C or she is practically irrational.<sup>20</sup>

Smith calls this 'the practicality requirement on moral judgment' – or just 'the practicality requirement' for short – and asserts it as a conceptual truth. Just from this claim, it is evident that Smith takes the necessary connection between moral judgment and motivation to also be a defeasible one. The necessary connection can be broken if an agent is 'practically irrational'. This might elicit criticism that Smith's view is too weak and that it could collapse into externalism. The reason Smith's view is nonetheless an internalist one is due to what is entailed by his position regarding practical (ir)rationality.

Smith endorses a view he calls 'rationalism as a conceptual claim': "the claim that our concept of a moral requirement is the concept of a reason for action; a requirement of rationality or reason."<sup>21</sup> Following an analysis of normative reasons, Smith further argues that when an agent acknowledges a reason for action, she shows most rationality by being motivated to act upon that reason. In other words, making a moral judgment is the result of deliberation about what it is morally desirable to do; and when we have a moral belief, we show most rationality if we have a desire to act upon it. Therefore it follows that if an agent is rational and acknowledges a moral requirement (a reason for action) she ought to be motivated to act upon it. Consider the following passage from Smith:

---

<sup>20</sup> Smith 1994: 61.

<sup>21</sup> Ibid: 64. 'Rationalism as a conceptual claim' should be understood as "the best analysis of moral terms." (65) Just as with his practicality requirement, Smith provides a formulation for this claim: "If it is right for agents to  $\Phi$  in circumstances C, then there is a reason for those agents to  $\Phi$  in C." (62)



“It is a platitude that an agent has a reason to act in a certain way just in case she would be motivated to act in that way if she were rational... And it is a consequence of this platitude that an agent who judges herself to have a reason to act in a certain way – who judges that she would be so motivated if she were rational – is practically irrational if she is not motivated to act accordingly. For if she is not motivated accordingly then she fails to be rational by her own lights.”<sup>22</sup>

Smith equates ‘judges herself to have a reason to act in a certain way’ with ‘judges that she would be so motivated if she were rational’. At first glance, these two ideas do not seem to be necessarily equivalent. However, Smith makes this claim on *reasons for action*, or normative reasons, by providing the following explanation. Smith observes that the examination into normative reasons must start with the odd fact that an agent’s deliberations on normative reason claims sometimes influences her actions, while sometimes they do not: “that accepting normative reasons claims can both be *bound up with* having desires and yet *come apart from* having desires.”<sup>23</sup>

So begin with the supposition that ‘believing that one has a normative reason x’ is equivalent to ‘valuing x’.<sup>24</sup> The question would then be what is the relationship between ‘valuing’ and ‘desiring’ (where desiring here is understood as the motivational state)? Often times these two concepts go hand in hand as, for example, people that value a good relationship with their parents also desire a good relationship with their parents. However, they are not identical states as one can also value something but not desire it and desire something but not value it. One can value exercise for good health without having a desire to actually exercise, as well as having a desire for cigarettes without valuing smoking. With these considerations playing a major role in his analysis of valuing, Smith concludes that ‘valuing something’, or the mental state of valuing (that is, to accept that we

---

<sup>22</sup> Ibid: 62.

<sup>23</sup> Ibid: 136.

<sup>24</sup> Ibid. 132.

have a normative reason to do something), must be a belief-like cognitive state rather than a desire-like non-cognitive state because no analysis of valuing as desiring can overcome the notion that we can also value something but not desire it.

Smith notes that a point which emerged from this analysis of valuing possibly as a form of desiring is that “though a rational agent desires... in accordance with the normative reason claims she accepts, an irrational agent may desire otherwise,”<sup>25</sup> and he goes on to assert that this notion suggests that normative reasons are subject to the following constraint.

C2     If an agent believes that she has a normative reason to  $\Phi$ , then she rationally should desire to  $\Phi$ .<sup>26</sup>

Smith believes this to be an accurate depiction of the conceptual connection between our moral beliefs and desires and supports this view by noting that it is a platitude to say that for any agent, “what we have normative reason to do is what we would desire that we do if we were fully rational.”<sup>27</sup> Smith believes this adheres to many of the platitudes regarding our conceptions of what is entailed by reasons for action and motivation.

Furthermore, throughout his book Smith mentions weakness of will, apathy, accidie and the like as paradigm cases of practical irrationality.<sup>28</sup> Regarding these forms of practical irrationality, with phenomena such as weakness of will and apathy, the agent seems to be in some sense ‘willingly’ going against her better judgment. That is, for Smith, the reason we can consider such phenomena as paradigm cases of practical irrationality is because the psychology of such an agent who is not motivated to pursue the content of her own genuine moral judgment clearly exhibits an incoherence in that she fails to desire (and so act upon) those very things she herself judges to be

---

<sup>25</sup> Ibid. 143.

<sup>26</sup> Ibid. 148.

<sup>27</sup> Ibid. 150. Smith cites Korsgaard (1986) to reinforce this notion as being a platitude.

<sup>28</sup> Ibid: 61, 120, 135.

the right or morally required thing to do, and where there is no ‘external’<sup>29</sup> interference to stop the agent from desiring accordingly.

Now contrast these cases to another type which Smith would also label as cases of irrationality – the agent who cannot be motivated to pursue her moral convictions either due to neurological damage or some form of a pathological condition. Unlike the previously mentioned phenomena of weakness of will or apathy – where the agent is in some sense willingly going against her better judgment – in the cases of neurological damage or having some sort of pathological condition, even Smith would grant that the agent here is not willingly failing to endorse the moral demands she has acknowledged in the same sense. Nevertheless, Smith would regard them as irrational in the sense that “the psychological processes of the perfectly rational agent are subject to no such interference.”<sup>30</sup> Therefore, for Smith, it is true that a rational agent must desire in accordance with the normative reasons she believes she has, and any case where that is not so should rightfully be labeled as a case of practical irrationality.

Going further into Smith’s internalist view, I mentioned that Smith also defends moral cognitivism.<sup>31</sup> This implies that he has a certain etiology in mind regarding the nature of the conceptual connection between moral judgment and motivation as well as a specific notion regarding what it means to judge that  $\Phi$ -ing is right and why therefore it is practically irrational to not be motivated thereafter. So if we were to reformulate Smith’s internalist claim in a more comprehensive manner, it might look something like the following:

Necessarily, if an agent judges that it is right for her to  $\Phi$  in circumstances C and is motivated to  $\Phi$  in C, then it is the content of the agent’s judgment that  $\Phi$ -ing is

---

<sup>29</sup> By ‘external’ here I do not mean literally external to an agent’s physical self or pertaining to her surroundings or environment. Rather, I mean ‘external’ more in the sense that it is not a natural part of what constitutes the agent’s will.

<sup>30</sup> Smith 1989: 95.

<sup>31</sup> Chapter 2 of Smith’s *The Moral Problem* is dedicated to arguing against non-cognitivism.

right in C itself that generates the motivation to  $\Phi$  in C; and if the agent fails to be motivated by her judgment that  $\Phi$ -ing is right in C, then she is practically irrational, because to judge that  $\Phi$ -ing is right in C is to judge that there is a reason to  $\Phi$  in C.

This schema captures all the major elements in Smith's internalist view: from cognitivism, to his modal claim, to his claim regarding the etiology of motivation, and to his thesis regarding practical rationality.<sup>32</sup> To summarize then, the reason Smith is an internalist, and the reason he supports some variation of the inverted commas objection, is because for Smith having a moral judgment is coming to the belief that it would be desirable to do that which is morally required. This is simply a matter of deliberating upon what it is that we are rationally justified in doing. Therefore, we are rational insofar as we endorse the moral judgment; and inversely, we demonstrate incoherence by neglecting the normative content of our moral judgments. So as far as rational agents are concerned, internalism is correct.

One final element that is worth mentioning about Smith's internalist view is his depiction of the 'good and strong-willed person'. In good and strong-willed persons, believing that doing  $\Phi$  is right, results in a desire (and thus motivation) to  $\Phi$ . The amoralist who is unmotivated by her sincere moral judgment deserves rational criticism because she fails to reflect a good and strong-willed person.<sup>33</sup> What does this rational criticism amount to? Part of the answer, I believe, is reflected in Smith's variation of the inverted commas objection against externalism. I discuss Smith's version of the inverted commas objection in chapter IV. In the rest of this chapter, I look to

---

<sup>32</sup> A more detailed discussion and explanation of what these terms refer to is provided in section 3.1.

<sup>33</sup> I question whether an amoralist really does deserve *rational* criticism, since I believe it is an open question whether the concept of practical rationality requires a necessary connection to morality. If instead, we were to take a line similar to the upcoming one I bring up when discussing Foot's view; that is, if the amoralist agent was not motivated due to the fact that she understood the moral requirement to hinder her personal goals, then the type of criticism the amoralist deserves might not be of the rational variety, but perhaps simply of a moral one.

defend externalism against the internalist claims laid out by Smith before I present my positive account of the (contingent) relationship between moral judgment and motivation.

## 2.2. On the Relationship Between Rationality and Motivation

R. Jay Wallace states that the interaction between (normative) reasons and motivation is such that reasons on their own do not have some ‘magically intrinsic’ motivational power to them, but rather that a further argument must be given. He cites the example that contributing to one’s pension fund in order to enhance financial security in old age is, in itself, merely an interesting fact about that course of action, but that it is perfectly possible for a rational agent to acknowledge that this fact obtains and still lack any motivation.<sup>34</sup> Wallace further asserts that the term ‘irrationality’ is best reserved to “refer to cases in which people fail to act and think in accordance with the normative judgment about their reasons that they themselves endorse.”<sup>35</sup>

But it seems entirely possible that people can fail to endorse the normative content of their reasons, let alone the fact that people can fail to act accordingly. For when an agent acknowledges that it is morally right to  $\phi$ , it seems this acknowledgment is another way of saying that, from a *moral perspective*, she has a reason to  $\phi$ . But this does not mean that if the agent is rational, she is therefore necessarily acknowledging from a more practical *commitment-to-action* sense, that she has a reason to  $\phi$ . That is, it seems fully plausible that rational agents sometimes do not take their commitments as far as *actually carrying out the action at stake*, even on the basis of the (necessary) moral acknowledgment. Of course in many agents, not just those who we call virtuous, their moral perspectives will align with their practical side so as to affect their daily practical lives in a direct sense. However, it is not a logical conclusion that these two viewpoints will necessarily match in every (rational) agent.

---

<sup>34</sup> Wallace 2006: 186-187.

<sup>35</sup> Ibid.

Furthermore, even if accepting a normative reason to  $\phi$  is equivalent to accepting that it would be desirable to  $\phi$ , it does not necessarily follow that, if an agent is rational, she will be motivated by that which she has acknowledged as desirable to do. We can see this from a non-moral perspective. If a person made the judgment that it would be desirable to run a marathon (or rob a bank), this does not mean, were she a rational agent, that she is therefore necessarily motivated to run a marathon (or rob a bank). Thus I do not think that we have reason to accept Smith's claim that 'a rational agent desires... in accordance with the normative reason claims she accepts', nor that such acknowledgment will necessarily lead to motivation.

Consider the following passage by Gary Watson:

Notice that, when things go as intended, practical deliberation involves making up my mind *twice*. Making up my mind about what is best to do is coming to a judgment: deciding *that* such and such is the thing to do. Making up my mind about what to do is forming an intention: deciding *to* do such and such... Although they typically coincide, these are importantly distinct forms of commitment. The distinction manifests itself in at least two familiar ways: first, when we fail to reach a decision about what is best to do because the reasons are unclear or indeterminate but we still must decide to do *x* or to do *y*; and, secondly, when we fail to follow our decision about what is best.<sup>36</sup>

There is correlation between this passage and Smith's cognitivist account in that Smith accepts that moral judgments are practical judgments in the sense that they involve practical deliberation. However, notice that in the passage, Watson asserts that an agent can 'fail to follow [on her] decision about what is best'. This claim, I believe, is not controversial because Watson makes a distinction which I do not think Smith makes. Watson makes a distinction within practical

---

<sup>36</sup> Watson 2003: 176.

judgments as it relates to motivation in that there are some types that are not related to motivation while others are. More specifically, judgments about what ought to be done in (moral) situations are not the same as ‘making up my mind about what to do... forming an intention’. It is the latter that has an intrinsic connection to motivation, whereas the former necessarily does not. As Elinor Mason put it, “There is an internal connection between motivation and ‘judgement’... if by judgement we mean something like decision or intention: Watson’s second sense. However there is no internal connection between thinking an act best and motivation.”<sup>37</sup>

Mason further elaborates that this distinction can be understood as being the difference between *seeing* the appropriate means to one’s ends, versus *pursuing* the best means to one’s ends<sup>38</sup>; where ‘pursuing’ of course does not imply that action necessarily has to follow, but is rather forming an intention towards whatever aims have been concluded from the deliberative processes that have taken place. I believe that Smith’s view trades-off on the ambiguity between the two forms. But following the distinction explained by Watson and Mason, charging an agent as practically irrational for failing to be motivated due to not being able to grasp, or *see*, the appropriate means is not an indictment against the agent’s *motivational sensitivity*, but rather a charge of a lack of *moral sensibility*. If this were the charge, it would make (weak) motivational internalism entirely trivial: for “the defeasibility of motivation must not undermine the claim that there is a necessary connection between [moral] belief and motivation.”<sup>39</sup> The charge of practical irrationality must indicate some other type of fault in the unmotivated agent, reflective of a more general impairment than simply having committed some error of ‘moral badness or moral weakness’.

---

<sup>37</sup> Mason 2008: 152.

<sup>38</sup> Ibid: 147.

<sup>39</sup> Ibid: 145.

I believe that Philippa Foot was arguing along these lines in her paper, *Morality as a System of Hypothetical Imperatives*.<sup>40</sup> As Foot notes, a rational amoralist who is not motivated by her moral judgment can only be considered as acting wrong in a moral sense, not in a rational sense.

The fact is that the man who rejects morality because he sees no reason to obey its rules can be convicted of villainy but not of inconsistency. Nor will his action necessarily be irrational. Irrational actions are those in which a man in some way defeats his own purposes, doing what is calculated to be disadvantageous or to frustrate his ends. Immorality does not *necessarily* involve any such thing.<sup>41</sup>

For Foot, practical irrationality involves ‘defeating one’s own purposes’ and ‘frustrating one’s ends’. In other words, practical rationality is a prudential notion that centers on an agent’s self-interests. Foot herself eventually ends up renouncing this view, calling it a ‘bad mistake’.<sup>42</sup> However, I believe the core idea of her view is still relevantly insightful and can be prompted as a prominent argument with a little supplementation. For the strength of Foot’s argument does not rely upon whether it is as a matter of fact truly the case that practical rationality involves a strictly prudential (or even just an agent-centered, as opposed to other-regarding) conception. Rather, it simply requires possible conceptions of practical reason that might not align with morality in the sense that it is possible to imagine that moral requirements might not have absolute rational authority; or as David Brink put it, “questioning the rational authority of moral reasons.”<sup>43</sup> That is, as long as there are such possibilities, with Foot’s claim exemplifying the groundworks, then it seems that we must at the very least recognize the possibility that the rational authority of morality should be treated as an open question.

---

<sup>40</sup> Foot 1972.

<sup>41</sup> Ibid: 310.

<sup>42</sup> Foot wrote a one page ‘Recantation’ (322) in *Moral Discourse and Practice* (1997), edited by Darwall, Gibbard and Railton. For an extended explanation of why she changed her mind regarding this view she had originally forward, see Foot chapter 4 (2001).

<sup>43</sup> Brink 2008: 292.



Note that there seems to be a sense in which the statement ‘there is always a reason to do as morality demands’ is a conceptual truth that everyone accepts. This is not a rebuttal to the arguments laid out so far, as that statement simply understands every moral obligation as entailing a moral reason, just like every legal obligation providing a legal reason, or every obligation of etiquette entailing an etiquette reason. Therefore, even if the internalist says that the moral reason provided by the moral obligation (in this sense) supplies the answer to those who question the grounds of obedience, this is not the kind of answer that would satisfy a rational amoralist’s query into practical (moral) reasons. For such an agent could question what justification there is to follow *those* moral reasons; and then what justification there is to follow the reasons offered for following the aforementioned reasons, and so on *ad infinitum*.<sup>44</sup> Or as Wallace additionally considers, perhaps she might even simply accept the truth of moral claims without accepting that they apply to her (or any agent for that matter) in any normatively significant sense.<sup>45</sup>

The key point is that what matters is what beliefs an agent could hold, not whether those beliefs are true or not. If an agent could hold conceptions of practical reason that might not align with morality in the sense that moral requirements might not have absolute rational authority, perhaps due to certain beliefs about morality and rationality that diverge, then it is conceivable that such moments of doubt (even if they hypothetically rest on substantive mistakes regarding the rational authority of morality) could lead to failure in motivation from her moral judgments. Of course, it would be much less problematic if the demands of morality and that of practical rationality always align. But undoubtedly an agent can, at the very least, perceive that the demands of morality and practical rationality do not always converge. So even if her lack of motivation is due to a mistaken conception of morality or rationality, the (principled) amoralist is conceivable.

---

<sup>44</sup> Shafer-Landau 2003: 166.

<sup>45</sup> Wallace 2006: 190.

### 2.3. The Externalist Claim

A way to understand the externalist claim is that the connection between moral judgment and motivation is contingent on: (i) a person's psychological make-up and (ii) on the perceived content of moral demands.<sup>46</sup> To say that the connection between moral judgment and motivation is contingent on a person's psychological make-up is to mean that there could be agents whose possession of certain mental conditions (such as depression, apathy, lethargy, etc.) or certain events in an agent's life that could alter her disposition (perhaps due to psychological trauma experienced by the agent during the event) can lead to a disconnect between moral judgment and motivation. However, this claim only seems to affect strong forms of internalism; as weaker versions can accommodate this claim by providing an understanding of the concept of a 'rational agent' as excluding these types of instances, in addition to false beliefs or improper reasoning, as forms of irrationality. Since the view that needs to be dealt with in this thesis is the weak internalist view, this sort of claim for externalism would therefore not be a very successful one.

The externalist claim that the connection between moral judgment and motivation is contingent on the agent's perceived content of moral demands includes at least two possible reasons for why there could be a disconnect. The first possibility would be that the agent perceives the content of the moral demand to be so impossible to achieve, so overly demanding, or so out of touch with perceived reality, that she cannot find herself motivated to follow it. That it would either be futile or practically irrational, in her mind, to even attempt to follow it; and this apparent sense of futility, impracticality or irrationality could then conceivably strip an agent of motivation. To elaborate on this, consider when an agent makes a moral judgment in which she believes that  $\Phi$  is right but also that the goal of  $\Phi$  is impossible to achieve. This seems to be conceptually possible,

---

<sup>46</sup> Shafer-Landau 2003: 151.

especially since there are certain normative ethical views<sup>47</sup> which have been criticized to be so demanding that in certain cases – perhaps because there are certain conditions that just cannot be met or certain circumstances that provide to be too great of an obstacle to the agent – they just do not seem to be satisfiable. In this type of situation, it seems conceivable that a feeling of futility which accompanies the moral judgment might stifle any sort of motivation that could or should have come with the moral judgment.

One important point to note though is that this argument only works if the agent does not disavow the value of the normativity that is contained in the moral content due to the impossibility (perceived or otherwise) of following the moral demand. In other words, there is the objection that this example faces the ought-implies-can edict; or more specifically speaking, the edict that what is impossible cannot be a moral obligation. However, whether that view is correct or not, what matters is how an agent perceives the moral demand in the judgment. As far as that is concerned, it seems entirely possible that an agent either does not consider such an ought-implies-can edict or personally holds a rejection of such a view.

The internalist here might respond that ‘thinking  $\Phi$  to be obligatory yet impossible at the same time’ – that is, implicitly rejecting the ought-implies-can edict – is simply a matter of irrationality. That one cannot rationally think she is morally obligated to do that which she finds impossible to do. But far from being the case that the ought-implies-can debate has been resolved, it would seem to me that arguing for irrationality based on the rejection of this *deontic* principle is basically presupposing a connection between rationality and ethics that needs to be argued for separately. Furthermore, even if one provides a valid conceptual argument for the connection, it is another thing entirely to thus conclude that the psychology of all moral agents fits neatly according to such a conceptual standard of rationality. For when the point regarding some agents seemingly

---

<sup>47</sup> Certain varieties of strong consequentialist and utilitarian views could serve as examples here.

rejecting the ought-implies-can edict is elaborated on, it is that their rejection of this view is not based on some clear conceptual foundation, but because their psychology is such that their implicit rejection of this view can be seen empirically by their judgments or actions. Let me provide a non-moral case as an example to support this view.

Imagine as a high school senior your parents implore you to apply to Harvard and Yale and you know that applying to such Ivy League schools is a worthy endeavor even if your parents did not make this request. However, you also know for certain that your grades and resume are just not good enough to get you in and that your parents just have an unreasonably rosy perception of your academics. Thus you feel that spending the time to apply to the two schools would just be a huge waste of time, followed by the obvious disappointment of the rejection letters. Nonetheless, your valuation of the worth of what it means to earn a degree from those schools is not changed. In this case, the general consensus seems to be that a sincere prediction of futility will ensure that no motivation follows even while you maintain the value of the non-moral demand.<sup>48</sup> If so, and if many already accept that this is a perfectly plausible explanation for the failure of motivation to arise, then there needs to be a separate explanation regarding what is markedly different between moral and non-moral cases.

A second way to elaborate upon why an agent might perceive it irrational to follow a particular moral edict is if moral and prudential demands conflict and so the agent, in perfect cognition, readily dismisses the moral demand – perhaps the thought of giving up on the prudential demand causes too much fear, worry, concern, etc. This is based off a classic externalist notion that sometimes moral obligations can conflict with certain demands of prudence or self-interested desires. Regarding such a scenario I want to list out a few possible variances. First, the moral demand seems particularly inconsequential. Perhaps she thinks it is a moral demand that will do

---

<sup>48</sup> Of course if the agent still filled out the applications, the motivation behind such action can be explained by a desire to please one's parents or being motivated by a sense of hope, for example.

absolutely no harm if not followed, or if not followed on just this one occasion. Second, the prudential consideration is of a particularly invaluable or indispensable sort; one which, if had to be sacrificed, would take a great toll or cost on the agent. In another sense, perhaps forfeiting the prudential consideration, in order to follow the demands of the moral obligation, might just be too dreadful – it might cause great fear or anxiety in the agent if having to be sacrificed. When the ratios are on such opposite extremes of the spectrum, it seems that the agent might not just be past the point of wavering, but too gripped with fear or anxiety (or greed in the case of incurring a cost) that she is just unable to be motivated to follow the moral demand.

What is important here as well is that the agent still understands that the moral demand is a genuine reason. It is arguable that emotional elements such as fear or anxiety could momentarily ‘blind’ the agent, so to speak, so as to allow for a charge of irrationality. But I think it is entirely possible for an agent who is filled with fear to nonetheless maintain complete control over her cognitive faculties insofar as judging what is at stake; as this is reflected by an agent’s feeling of guilt or cowardice towards her own lack of motivation or inability to act upon her moral judgment. The fact of the matter is we often acknowledge normative reasons that are not necessarily in line with what we endeavor or want to do. Therefore, even if we think we have to do it, if the reason for doing it does not seem particularly appealing to us, why would it not be conceivable for us to lack the motivation to do it?

For instance, I acknowledge the normative significance of legal considerations, but I may fail to be motivated by a certain one simply because I do not hold the normativity of *that* legal consideration to sufficiently outweigh conflicting reasons or potential costs. Take the example of a fairly recent law enacted in South Korea that makes it illegal for private teaching institutions (known as ‘hagwons’) to be open later than 11 p.m. and let’s suppose that I am a Korean parent in South Korea. I may, for instance, believe that although I cannot dispute any fine levied unto me if I

were to transgress because I do acknowledge the normative authority of the law, nonetheless I could have a certain view on education and child rearing that inclines me to send my child to study at such an institution past the allotted time if that academy is willing to protect my child from any legal dangers and so forth. Perhaps this reveals my limits as a law-abiding citizen, but does it indicate a failure in rationality if I am not motivated appropriately from my acknowledgment that this law is normatively significant? I find it to be far from obvious that this is the case.

The internalist could argue whether the example is applicable – that the domain of laws and that of morality are sufficiently distinct so that this example is inapplicable. That perhaps the normativity entailed from legal requirements is somehow weaker, or at least different, than that of moral requirements. However, I think there is good reason to consider the arguments of this example as applicable since the basic structure of the normativity of legal considerations – that it applies to everyone involved and that it makes no exceptions based on an agent’s preferences or desires – seem to be no different from that of the normativity of moral considerations. Wallace, who is an internalist himself, concedes that “it is not irrational to fail to be motivated to act on reasons that one acknowledges if those reasons are either defeated themselves, or such as to render an action one of a plurality of eligible or valuable options for choice,”<sup>49</sup> and further adds that “it would be implausibly strict.. requiring that agents should have an occurrent desire for every normative reason they acknowledge to obtain... Rational agents will be motivated to act on the reasons they acknowledge to obtain, so long as they do not take those reasons to be defeated by other normative considerations, and so long as they have not resolved to pursue other eligible options.”<sup>50</sup> But even most internalists (at least the ones which comprise the target of my thesis)

---

<sup>49</sup> Wallace 2006: 188.

<sup>50</sup> Ibid: 187.

concede that moral reasons are not “*always* so compelling that they defeat whatever reasons might count against the actions morality demands.”<sup>51</sup>

#### 2.4. Action-Guiding Not the Same as Motivating

As a final point in concluding my arguments against Smith’s internalist contention, I think we have little reason to accept the assertion of a necessary connection between moral judgment and motivation because I believe what drives the internalist sentiment is actually a conflation between what it means for something to be normative and for that thing to be motivating. Wallace argues that part of the conception of moral judgments as being normative entails that such acknowledgment by an agent has to be followed by motivation. This is necessary in order to show that “those who are morally motivated are somehow responding appropriately or correctly to the moral distinctions they grasp.”<sup>52</sup> However, far from being an obvious truth that there is such a motivation-entailing normative requirement (as externalists deny that there is such a dimension to capture<sup>53</sup>), I think that Wallace’s claim is an example of precisely this conflation, and so I look to explain my concern against internalism by arguing against a direct and necessary connection between the normativity of morality and moral motivation.

First let me loosely define something being normative as making a claim on how something ‘ought’ to be, whether something is ‘good’ or ‘bad’, or whether some action is ‘right’ or ‘wrong’. Needless to say then, moral judgments, the content of which generally contain terms such as ‘ought’, ‘good’ or ‘right’ are obviously normative in nature. However, as far as moral judgments relating to motivation goes, it is important to note that it is not the normative *fact* simpliciter, but rather the effect of the normative *force* of these facts on the agent making the judgment, that determines

---

<sup>51</sup> Ibid: 188.

<sup>52</sup> Ibid: 194.

<sup>53</sup> See Svavarsdóttir 2006: 172 and Zangwill 2003: 150.

whether she is motivated from her moral belief or not. Whether or not this is obvious, there seems to be a general conflation of these ideas. The fact of the matter is, normative content (or what I referred to as the normative fact) is strictly reflective of the ‘action-guidingness’ of morality. Furthermore, because this action-guidingness of morality does not have a connection to motivation, it is thus a misattribution to say that the normative nature of moral judgments suggests an entailment of motivation. That is, action-guidingness in morality is a distinct characteristic of morality that is outside the purview of the *psychological* domain that involves motivation.

The primary indication of a conflation comes from the widely accepted view that morality, in general, is supposed to be ‘practical’ or ‘action-guiding’ in an intrinsic sense. For instance, a moral principle that cannot show us what to do is often regarded as either not being a real moral principle, or trivial in the sense of being incapable of serving as a moral principle. Thus, the term ‘action-guiding’ in meta-ethics is commonly referred to as a natural characteristic inherent to morality in general. Brink explains this idea using the relation of the concepts ‘action-guiding’ and ‘practical’.

It has seemed to many people that moral considerations are practical in some very important sense. Agents engage in moral deliberation in order to decide what to do and give moral advice with the aim of influencing others’ conduct in certain ways. We regard moral considerations as important practical considerations. We expect people who accept moral claims to make moral judgments to act in certain ways.<sup>54</sup>

This kind of perception naturally lends favor to the internalist understanding of the relationship between moral judgment and motivation, as noted by Smith.

[M]oral judgments seem to be, or imply, opinions about the reasons we have for behaving in certain ways, and, other things being equal, having such opinions is a

---

<sup>54</sup> Brink 1986: 25.



matter of finding ourselves with a corresponding motivation to act. Let's call this the 'practicality of moral judgment'.<sup>55</sup>

However, the basic action-guiding conception of moral judgments as guides to conduct should not be construed as an argument for motivational internalism. That is because to say that someone *has* a moral obligation (that there is action-guidingness in the moral considerations) is not the same as saying that someone *feels* or *is moved by* the moral obligation. In other words, to say that someone has a moral obligation does not necessarily mean that an agent will be motivated to act as such just from that notion itself. Understanding a moral principle as a guide to conduct and an agent being motivated by the comprehension of that moral principle are two separate notions, so that even if we take it as a conceptual truth that moral judgments purport to guide action, that does not entail that they necessarily succeed in doing so.

Sigrun Svavarsdóttir provides two arguments for differentiating the realms of action-guidingness and motivation. For the first argument, she asks what it is to understand a piece of language as expressing a prescription, guidelines or directive.<sup>56</sup> She rejects any possible conclusion along the lines of such an understanding necessarily entailing motivation to undertake the prescribed action. She offers the following non-moral example for why simply understanding a moral idea fails to motivate, "I can understand a recipe [of how to broil salted cod] and pass it on to others without having any inclination to follow it." She next puts forth for consideration whether a stronger discernment of *accepting* a moral idea would be a kind of manifestation of motivation to follow its directions. She is quick to also deny such as being the case as she notes that she can both "accept the recipe as a guide to broiling salted cod, yet not be at all inclined to broil salted cod," hence not engaging in the project at all, or "accept the recipe as a guide to broiling salted cod and broil salted cod without being in the least motivated to follow the recipe" even if she finds nothing

---

<sup>55</sup> Smith 1994: 7.

<sup>56</sup> Svavarsdóttir 2006: 171.

wrong with the recipe itself but just simply because she may wish to do things differently from the way the recipe prescribes. “In other words, it is possible to engage in a project and accept something as a guide – even a good guide – to such a project, yet not be moved to follow its directions.”<sup>57</sup> What this means is that while adequate understanding of a moral judgment requires, at minimum, that the agent accepts that it is a moral guide to action (or living, on a broader interpretation), an adequate understanding of a moral judgment does not require motivation to follow.<sup>58</sup>

Svavarsdóttir’s second argument adds that any internalist defense for such a connection between the normative nature of moral guidelines and moral motivation is conceptually confused. For as she notes, “this defense relies on an implausible understanding of the claim that moral judgments are guides to action: the directional role essential to moral judgment cannot be reduced to a motivational role.”<sup>59</sup> In order to better understand what Svavarsdóttir is meaning here, consider the following idea. Take the moral principle ‘do not harass others’. If a person felt no motivation towards that principle and were to ignore her own moral judgment (to not harass others) and acts contrary to it instead, the ‘action-guidingness’ of that moral judgment does not disappear or get diminished in any sense due to the agent’s actual motivations or actions. Rather, the judgment is action-guiding nonetheless in the sense that it maintains a guideline which the person happened to ignore or transgress.

---

<sup>57</sup> Ibid.

<sup>58</sup> There is, however, an important disanalogy between the examples provided by Svavarsdóttir of guidelines for cooking to guidelines provided by moral judgments. It is of course that the latter are inescapable in that they are not dependent on the agent. However, it is important to note that this inescapability once again refers to the normative aspect – the action-guiding practicality of morality – not the psychological aspect of motivation. The guidelines of the cooking recipe are not binding on me to follow unless I wish to do so intentionally or have committed myself to do so otherwise. Therefore, I make no mistake in acting contrary to the cooking guidelines. Moral judgments do make such a normative claim regardless of whether I wish to follow it or not. However, as I have argued already regarding the conflation of normativity and motivation, the normative inescapability of moral judgments cannot be explained in terms of their actual or hypothetical motivational role.

<sup>59</sup> Svavarsdóttir 2006: 170.

Furthermore, if moral motivation were to have a direct link to the action-guidingness of moral judgments – where, for example, “the claim that a judgment guides conduct is that it motivates the agent to act in one way rather than another”<sup>60</sup> – then the *decisiveness* of any particular moral guideline would vary from case to case depending on the motivational impact of the judgment on the individual involved. However, that would seem plainly absurd. It is possible to consider the ‘weight’ of any given action-guiding prescription of a moral judgment to vary against competing moral prescriptions (à la W.D. Ross’ ‘prima facie duties’<sup>61</sup>) but it seems like a strange ethical theory to say that the weight of such prescriptions can vary depending on how much an agent is motivated by it. If these arguments seem plausible, I think we have good reason to be wary of the internalist assertion of a necessary connection between moral judgment and motivation as a conceptual truth.

### III. The Source of Moral Motivation

#### 3.1. Non-Humean Externalism

In the introductory section I stated that my position should be understood as taking an externalist stance, whereby rejecting Smith’s assumption two (motivational internalism). I further noted that, as an additional consequence to what my position amounts to, I reject assumption three (the Humean theory of motivation) as well. My rejection of these two views is another way of saying that I argue for the following two claims:

4. That moral motivation is not a *necessary* entailment from moral judgment, and;
5. That moral motivation does not involve a basic, underived desire-like non-cognitive state.

---

<sup>60</sup> Ibid: 170.

<sup>61</sup> Ross 1930.

In order to support the aforementioned claims, I argue for a position that I believe is hardly popular in this field. This is because the two claims, (4) motivational externalism and (5) non-Humeanism, seem to conflict with each other. Of course it is not uncommon to accept one or the other position and there are many philosophers who do so, but there do not seem to be many who accept both positions simultaneously. Let me briefly explain why someone might think holding both of these positions together might be conflicting and then I will follow with an explanation of why this is not the case.

First of all, claim (5) (the claim for the falsity of Humeanism), in conjunction with cognitivism, not only implies that moral judgments are to be understood as representing some type of belief-like cognitive state, but also argues for the possibility of a belief-like state being sufficient to motivate. As I am looking to assert in support of this claim then, moral beliefs alone are capable of motivating, thus are sufficient, if they motivate those who make corresponding moral judgments without having any antecedent desire-like state. However, if this is so, then it would seem that moral judgments, as beliefs, have to be necessarily motivating. Of course a sufficient condition does not imply a necessary condition, but the claim that moral judgments motivate seems to be the same as saying that moral judgments entail motivation, which would then seem to essentially be a reflection of the internalist view. That is, unlike the externalist that posits an external source as the cause for moral motivation, the internalist argues that moral motivation is essentially entailed from moral judgment. I believe this idea, although mistaken, is on the right path. However, internalism would then be contrary to my initial claim of externalism. So I offer the following explanation for how these two claims can be held simultaneously.

Moral judgment internalism is loosely understood as that moral judgment is *necessarily* motivating. From this, it seems unproblematic to say that moral motivation is *intrinsic* to moral judgment. I reject the former and accept the latter. If we consider internalism to have multiple

facets, we can define these two points distinctively.<sup>62</sup> (i) The first can be understood as being internalism's *modal claim*. It is by virtue of either being a conceptual or empirical necessity that moral judgments *necessarily* motivate. (ii) The second facet can be understood as being internalism's *etiological* claim of motivation. This is a claim regarding the *nature* of the connection that holds between moral judgments and motivation (which is used to provide an explanation for how moral judgments *cause* motivation). Cognitivists and non-cognitivists have decidedly differing explanations regarding the nature of the connection which I discuss in more detail towards the latter half of this section. For now, when considering the two facets here (in addition to the fact that I am more sympathetic to the externalist view), I believe that the *necessary* (modal) connection is the crucial element depicting internalism. Therefore insofar as rejecting the necessary connection, I take my view to be externalist.<sup>63</sup>

By rejecting the Humean theory of motivation and asserting that moral judgments are simply expressions of an agent's beliefs, my view can be understood as that moral beliefs are the mental states responsible for moral motivation and that such motivation is intrinsic to that very moral judgment. But as I contend, a moral belief may be *intrinsically* motivating without being *necessarily* motivating. Since I assert that internalism (of the cognitive variety) is a view foremost regarding a necessary connection between moral beliefs and motivation, I do not believe that it is contradictory to hold the position I am trying to defend.<sup>64</sup> We can see this point more clearly with the following two statements:

---

<sup>62</sup> Cueno 1999: 361.

<sup>63</sup> I am open to the idea that both elements are required to be an internalist or both elements have to be rejected in order to be an externalist. But I do not think that this is an important point as far as clarifying my position is concerned, so I will ignore this potential issue in my thesis.

<sup>64</sup> This position is not original to me as Russ Shafer-Landau argues for a similar position. See Shafer-Landau 2003: 147.

6. *All* moral judgments motivate agents that make or hold such judgments. Thus moral judgments necessarily motivate at all times.
7. *When* moral judgments motivate it is necessarily due to something internal to the moral judgment.

Statement (6) does not explicitly state anything descriptive regarding the source of motivation, but simply states the fact that there is a consistent and necessary connection between moral judgment and motivation. Therefore, this statement is neutral regarding cognitivism/non-cognitivism or Humeanism/non-Humeanism. It could be the case that the necessary connection between moral judgments and motivation is due to the fact that moral judgments simply are expressions of one's conative state. It could also be the case that moral judgments are actually fact-stating but that desires are still always involved in some sense which leads to the motivation. The possibility that moral judgments are purely belief-like states that somehow motivate is also available.

Statement (7) explicitly makes a reference to the source of motivation by asserting that the motivation that follows a moral judgment is due to something *internal* or *intrinsic* to the moral judgment. While the statement is vague as to the precise description of the source that composes a moral judgment, the idea is that motivation occurs in virtue of the very nature or content of the moral judgment. We can consider some common views that reflect this position: such as the agent's acknowledgment of the normative significance of the moral considerations being sufficient to motivate; or that such acknowledgment generates the relative desire-like state needed for motivation; or something along the lines of describing the psychological structure of an agent as being influenced by what is pertained in the semantic content of the moral judgment once the agent gains cognitive access to it. However, (7) is quiet on the degree of efficacy of motivation from moral judgments. As I have stated above, I consider statement (6) to be the depiction of motivational internalism proper.

There might, nevertheless, be an objection that statement (7) would seem to entail that every time a moral judgment is made, motivation would follow. It would seem that if motivation is intrinsic to moral judgment, it would be the case that necessarily motivation has to be entailed from a moral judgment. I think this misses the point of the distinction I have tried to make and brings us to the explanation of internalism's etiological claim of motivation. When we say 'motivation is entailed from a moral judgment' what is important is how we interpret the concept 'entailed' with respect to an account of the etiology of moral motivation. On the one hand, entailed could mean that motivation is *built into* the moral judgment. We can call this a *constitutional* claim – that motivation is somehow part of the constitution of a moral judgment. Non-cognitivists would maintain this claim and assert something along the lines of 'a moral judgment that  $\Phi$ -ing is right is *identical* to the agent's motivation to  $\Phi$ '. On the other hand, entailed could mean that motivation comes into existence (even to the slightest degree) whenever a moral judgment is made by an agent. This way of understanding entailment is that motivation is *generated* from one or more elements that compose a moral judgment but is not a constitutive part of the moral judgment itself. This understanding generally reflects the cognitivist's account that 'the propositional content of a moral judgment that  $\Phi$ -ing is right *gives rise* to or *produces* the motivation to  $\Phi$ '.<sup>65</sup> For lack of a better term, let me call this a *causational* claim.

Now part of what it means to take my position is to be working under the view that the entailment of motivation is a causational claim, not a constitutional one. While this should be apparent since I accept cognitivism regarding moral judgments, this can also be shown from both the externalist and non-Humean perspective. For under externalism, any argument for the plausibility of an amoralist would only work as long as motivation is not built into the moral judgment. Since if motivation was built in as part of the structure of a moral judgment, it would

---

<sup>65</sup> Smith 1994: 72.

then be impossible for an agent that makes a moral judgment to fail to be motivated. But externalism is precisely the view that argues for such a possibility of failure to be motivated from a moral judgment. The same type of reasoning applies when considered from a non-Humean perspective. As mentioned, the non-Humean view can be understood as that an agent's moral beliefs sometimes generating some derived mental state (alike that of a desire) which is then what accounts for motivation. As the derived mental state is not a necessary feature of what makes a moral judgment – since under the non-Humean model, the moral judgment is some form of moral belief – the only logical way to understand the motivation entailment under this view is as a causal claim, rather than the constitutional one. Understood in this manner then, it seems clearly possible that, even though the source of moral motivation is internal to the moral judgment, there are a few ways to dissolve the necessary connection between them.

Russ Shafer-Landau offers one such way by suggesting the possibility of *motivational defeaters* extinguishing the motivation that would have come about from the moral judgment.<sup>66</sup> The addition of motivational defeaters provides an account for instances when such intrinsically motivating beliefs fail to exert motivation. Motivational defeaters can be competing beliefs, desires, or even non-mental factors such as physical duress or exhaustion, which influence the agent by severing the connection between moral judgment and motivation. Shafer-Landau offers up two ways we can understand the notion of motivational defeaters, depending on whether it is the case that any relevant possible motivation from a moral judgment is completely distinguished or simply subdued when an agent lacks feeling motivation from her moral judgment.

What he calls 'motivation *pro tanto*' is motivation that, when overridden by competing factors, nevertheless remain as remnants or even in such a small degree that it might not be readily

---

<sup>66</sup> Shafer-Landau 2003: 148.



noticeable.<sup>67</sup> Still, even if it were causally inefficacious, the motivation is not completely dissipated and this is enough to support certain weaker versions of internalism. The idea would be that when intrinsic motivation is understood as *pro tanto*, then any moral judgment that exerts motivation will always exert motivation no matter how small in degree. So, moral judgments that are intrinsically motivating will also be necessarily motivating.

Alternatively, Shafer-Landau calls the second variation ‘motivation *prima facie*’.<sup>68</sup> Motivation of this kind, when affected by defeaters, is completely extinguished rather than remaining as remnants. Since any possible motivation that might have come out of a moral judgment is completely non-existent, it is correct to say that there was no necessary motivation entailed from the moral judgment. Understanding motivational defeaters in this manner is one method the intrinsic but not necessary externalist view I am advocating can be shown as plausible.

However, I do not prefer to rely on this distinction between motivation *pro tanto* and *prima facie* to support my view because the utilization of motivational defeaters to show a possible disconnect seems to infer an all things considered much stronger connection between moral judgment and motivation than what I have in mind. That is, if motivational defeaters play the essential role in breaking up the necessary connection between moral judgment and motivation, then it would seem that, were there no such ‘hindrances’, an agent would necessarily be motivated after every moral judgment she makes. However, this then opens up the possibility that this version of externalism is not much different from particularly weak forms of internalism, and I wish to avoid the possibility of such a conclusion. Under the picture I look to draw, although moral motivation has its source within moral judgment, it is not necessarily part of the nature of a moral

---

<sup>67</sup> I do believe that although not *readily* noticeable, the motivation has to be able to be ‘sensed’ in some form if the agent searches for it through introspection. If this condition is not necessary, then I do not see how even the weak internalist can maintain a convincing argument for the existence of said motivation without begging the question.

<sup>68</sup> Shafer-Landau uses the *pro tanto/prima facie* terminology following an analogous debate on the nature of practical reason and intrinsic value between Moore and Ross.

judgment to be motivating to an agent who acknowledges its normative significance. Thus, what I believe to be a more appealing view is that perhaps there is some further step required between that source and motivation for the agent to be motivated by her moral judgment. In the remaining parts of this chapter then, I look to provide an explanation of this ‘further step’ that seems to be needed between moral judgment and motivation which can adequately demonstrate an intrinsic yet contingent connection.

### 3.2. A Starting Distinction Between Two Types of Moral Judgments

The position I am advocating is that although moral judgments are responsible for motivation, it is not the case that moral judgments necessarily motivate.<sup>69</sup> This naturally favors a cognitivist reading of moral judgments, since if moral judgments were composed of some desire-like non-cognitive state, it would be difficult to see how they do not necessarily motivate. I take this as the starting point of my positive account and look to explicate upon this view. I begin from the cognitivist premise that moral judgments purport to be descriptive in the sense of aiming to portray some aspect of the world as being true. If moral judgments purport to be descriptive of some moral facts of the world, then one of the consequences of making a moral judgment must be providing an account or description of some objective idea.

Recall the quote by Watson. I mentioned that Watson makes a distinction within practical judgments as it relates to motivation which Smith does not make. While Watson’s terminology distinguished between *judgment* (about what ought to be done) and *intention* (about doing that which ought to be done), I think this is another way of simply stating that within the complexities of the human mind, where even one single moral judgment can contain several steps of the reasoning

---

<sup>69</sup> Of course this does not exclude the possibility that moral beliefs *usually* motivate. Since my view argues for a contingent connection, it need not be the case that there is a general lack of motivation following a moral judgment, only that it is not a necessary connection.

process, it is equally plausible that there are multiple judgments aptly fitting under a wider umbrella moral judgment which are individually sufficient to be moral judgments on their own, but can also be part of the two step process as Watson illustrates. This opens up the possibility that an agent can hold a moral belief in more than just one way.

Consider the moral judgments (i) ‘donating to charity is right’ and (ii) ‘I/one ought to donate to charity’. I argue that whereas the former should not be linked to motivation, strictly speaking, the latter can. Let me begin by clarifying the distinction I am looking to make between the two judgments. First, the difference between the two judgments is not one between evaluative versus normative judgments. As both judgments contain normatively loaded concepts (‘right’ and ‘ought’), they both appropriately belong in the category of being normative judgments. Secondly, the distinction is not one of content as both judgments are concerning the same content – the moral normativity of donating to charity. However, even from within the category of normative judgments, there seems to be a specific difference between the two propositions in question: a difference in how an agent might come to see what the judgment is trying to assert. In the case of moral judgments of the form ‘ $\Phi$  is right’ and ‘I/one ought to  $\Phi$ ’, the content of the first judgment can be seen as more aptly describing how the world is (or should be), whereas the content of the second judgment explicitly adds how agents should *act* regarding the subject matter. Therefore, it seems that we can distinguish the two moral judgments in the sense that an agent might *perceive* each normative judgment differently. The sense I am trying to convey is that an agent might perceive the first judgment as a descriptive assertion of the way a certain moral aspect of the world is or should be, whereas the second judgment might be perceived more precisely as a *moral imperative* due to its prescriptive characteristics. To clarify further, the difference in the agent’s perspective of the two judgments should be understood as a psychologically affecting distinction.

Nonetheless it seems that the two judgments are so tightly connected that it is logically necessary that the first judgment would entail the second. After all, if an agent asserts that it is morally right to donate to charity, and we assume that a rational agent also acknowledges that a feature of morality is that it applies universally to everyone in inferring under what circumstances and to whom this assertion applies to<sup>70</sup>, the rational agent would naturally include herself as part of the multitude of agents to whom this assertion applies. If she therefore does not conclude that she ought to donate to charity from simply her moral judgment that donating to charity is right – that her moral judgment that donating to charity is right *gives her a reason* to donate – we might say that she must be irrational. I think this is only half right. I would also accept the logical connection between the two judgments. However, I do not think the connection is quite as tight as it might seem from a psychological perspective.

Derek Parfit states that “while reasons are provided by the facts, the rationality of our desires and acts depends instead on what we believe.”<sup>71</sup> He further asserts that reason is the more fundamental idea of the two; that is, being rational ‘should’ result in responding to reasons. In the quote, Parfit was noting a difference between *reasons* and *rationality*. I think this difference is often overlooked since the two terms are often considered to be so closely related. As Parfit’s quote indicates, the difference is one between fact and mental state. Therefore, following this distinction by Parfit, insofar as motivation is concerned, whereas the former might only be perceived by the agent as involving judging what is right (therefore the judgment simply offering some fact to the agent), the latter seems to involve judging whether we ought to do that which has been asserted by the first judgment (in a sense, directing the agent as to how to proceed in her behavior). Going back to Watson again, his passage lent support for an understanding of a distinction between two types of

---

<sup>70</sup> Moral relativists would deny this. I leave this issue aside and will just assume an objective understanding of morality.

<sup>71</sup> Parfit 1997: 99.

(moral) judgments and that the connection between them can come apart (at least psychologically speaking) as his statement that ‘these are importantly distinct forms of *commitment*’. I emphasize ‘commitment’ because this is precisely what I will argue as being the difference between a moral judgment that motivates and one that does not.

### 3.3. Moore’s Paradox as an Analytical Tool

Above I tried to explain two ways an agent can hold moral beliefs. I believe that this might then be able to account for both amoralists and moral agents being able to make sincere moral judgments of the same kind. To elaborate, having such two ways of holding beliefs reflects the possibility that amoralists make moral judgments alike those of moral agents insofar as making an *epistemic commitment*, but without a further commitment in a motivating sense. In this section, I look to support my appeal of two ways of holding a moral belief using a moral variation of Moore’s paradox to serve as an analytical tool.

Consider the following assertion classically depicted to reflect Moore’s paradox:

‘It is raining but I don’t believe it is raining.’

The two conjuncts are logically consistent, and thus not an (obvious) contradiction. However, common sense tells us that it is nonetheless absurd to utter the two ideas at the same time in the sense that they cannot be both asserted as true. I believe that Moore’s paradox is helpful in my discussion because, in considering the paradoxical nature of the two utterances, it allows for the possibility of holding a belief in different ways which could reflect what is happening with amoralists and moral agents. I will suppose that this is the case but one might object that the analogy does not work because the absurdity arises from the fact that the utterance of the belief in the second conjunct (the ‘belief that it is not raining’) is used like an assertion and thus contradicts the assertion of the first conjunct (‘that it is raining’). That is, the absurdity occurs on the level of

assertions, not beliefs. However, I think that the absurdity is just as easily explainable on the level of beliefs in that the agent making the assertion ('that it is raining') *implies* that she believes that which she is asserting, and so it could actually be the beliefs that are contradictory along the lines of where the agent could be considered as implying:

'I believe it is raining, but I don't believe that it is raining.'

Using this form of Moore's paradox then, we can rewrite a moral variant in a manner such as:

'I believe that doing  $\Phi$  is right, but I don't believe I ought to do  $\Phi$ .'

Because the mystery of Moore's paradox is that the statement in its entirety can be true, logically consistent and not (obviously) contradictory, the amoralist agent should not have any problem holding (believing) both conjuncts of the statement at the same time.

However, the two conjuncts tell two separate stories reflecting the amoralist's mindset which can be explained by the notion, for example, that if she had a practically rational reason not to do  $\Phi$ , such as if doing  $\Phi$  would require a great sacrifice the amoralist agent is not willing to make. The first conjunct would represent an epistemic commitment made by the amoralist towards the normative content of the moral judgment, while the latter would represent the absence of a more embracing state of belief involving motivation. To make more clear what I have in mind, I suggest substituting the terms 'belief' in each part with terms that conceptually show a distinctive mental state but can also be (at least loosely) reflective of the concept of 'believing something'. I propose using the term 'acknowledge' for the term 'belief' in the first conjunct, and the term 'endorse' for the 'belief' term in the second conjunct. This modified (amoralist) version of the moral variant of Moore's paradox can then be shown as:

'I *acknowledge* that doing  $\Phi$  is right, but I don't *endorse* that I ought to do  $\Phi$ .'

Understood in this way, I believe we no longer have a case where the conjunction of the two expressions is absurd if uttered together. So for the sake of argument, if an amoralist can sincerely

make a moral judgment and not be moved, then I think it is fair to say that this variant of Moore's paradox should no longer be considered a paradox, but can be looked at as a solution instead; one which reflects the amoralist's mental state. As such the value of this Moorean 'solution' lies in the fact that it allows for the acceptance of forceful normative reasons, while at the same time being able to explain the amoralist's lack of motivation. The rational amoralist simply does not hold her belief in such a way that is associated with a necessary relationship to motivation.

### 3.4. The Operative State Within Moral Judgments Responsible for Motivation

Up to now, I have focused on describing the distinction of two ways of holding a moral belief, and arguing for the contingent connection between moral judgment and motivation. This was in effect, in order to reflect the mindset of an amoralist agent. But what about instances of when motivation does follow a moral judgment? I have yet to explain *how* motivation actually follows from a moral judgment when motivation is present. In order to understand this aspect of my account more clearly, it will be helpful to first discuss the non-Humean mechanism involved in the generation of motivation. While one of my main goals of this thesis is to focus on the internalist/externalist perspective, I feel that my positive account can be strengthened if I briefly discuss why I believe that understanding the relationship between moral judgment and motivation in this non-Humean sense is the most plausible account. That is, that moral belief somehow generates motivation when an agent has a deeper level of commitment than simply that of an epistemic one.

Part of what is considered unattractive about the externalist view is that the contingent nature of moral motivation seems to make an agent's motivated behavior somewhat arbitrary, as Wallace has noted:

There is nothing in this view to undergird the claim that those who are morally motivated are somehow responding appropriately or correctly to the moral distinctions they grasp... from the point of view of practical reason there is nothing to require that an agent who endorses moral claims should be motivated to comply with them. In this respect, the desire to be moral seems to be a mere optional extra, something some of us just happen to have – rather like a taste for clams or the color azure.<sup>72</sup>

The concern that is expressed by Wallace's passage is that positing this type of disconnect between moral judgment and motivation seems to distance the character of a morally motivated agent away from how we typically understand a 'virtuous agent' to be.<sup>73</sup> While I personally am not convinced that this is a serious problem, a more salient feature that I believe supports the non-Humean view is the following.

From an empirical perspective it is evident that the force of motivation, at least sometimes, necessarily correlates to the extent that an agent believes some object to be good, worthwhile, valuable, and so on. For example, the fact that I believe that, at this point in my life, academic success is more important than getting married and raising a family provides the explanation for why I am more motivated to spend my free time studying rather than going out and socializing. However, if several years down the road, I believe that it is now more important to get married and raise a family over achieving academic success, then this will, in the same way, explain why I am more motivated to mingle and meet people rather than work on my research. The way Philip Pettit explains the idea is that the degree of my motivation in relation to some x is my 'estimate of the

---

<sup>72</sup> Wallace 2006: 194.

<sup>73</sup> McDowell advocates for a purely cognitive view of moral motivation as it reflects the characteristic of the virtuous person's psychology who only needs to *perceive* what morality requires to be motivated to act. See McDowell (1978 & 1979).



objective goodness or utility' of that x.<sup>74</sup> That is, the presence of a certain motivation is entailed by the presence of a certain belief. Therefore, the sensitivity of the force of these motivations in relations to the beliefs which they mirror, show that these motivations are dependent on the correlating beliefs to the extent that the best explanation is that the motivations were produced by the beliefs. And this in turn supports the non-Humean mechanism that beliefs can generate certain motivational force which is what I mean by moral motivation being intrinsic to moral judgment. So with this understanding of how it is possible for moral beliefs to generate motivation, I think it will be easier to understand my account of what the operative state responsible for motivation is within such moral judgments.

Recall that one of the aspects of the distinction between moral judgments of the form 'doing  $\Phi$  is right' and 'I/one ought to  $\Phi$ ' is that whereas the former by itself lacks the necessary psychological force to motivate an agent, the latter also requires something extra, which is why I have argued for an intrinsic but contingent connection between moral judgment and motivation. Also recall statement (7) in section 3.1 where I explained what I meant by motivation being intrinsic but not necessary to moral judgment: *when* moral judgments motivate it is necessarily due to something internal to the moral judgment. In this statement, there is a further implication that I have yet to specify: that there is a subtle yet important difference between moral judgments where an agent is and is not motivated. Svavarsdóttir's recipe example invites the question of how motivation follows a moral judgment if it is the case that not even acceptance of a moral guideline suffices to motivate an agent. She offers the reason that there is a difference between accepting something as being *a* guide versus accepting something as *one's* guide: "Arguably, a person has not

---

<sup>74</sup> Pettit 1987: 531.

accepted something as *his* guide... unless he follows, by and large (barring weakness of will), its directions.”<sup>75</sup>

So consider the following. Even if we grant that we actually know very little as certain about what is going on in an agent who seems to be motivated by her moral judgment, introspection (or reflecting upon our experiences) seems to indicate that in order for an agent to act upon her moral considerations, she must have in some way experienced a *call to action*. What I suggest is that this experience is an agent essentially ordering herself to follow what the moral imperative is requiring of her. In other words, the agent essentially understands the moral imperative as a *personal imperative*. Another way to describe the phenomenology of this call to action is that for an agent to experience a personal imperative, it is for her to understand that her moral judgment is something for her to pursue. It is the experience of reflecting on one’s own moral judgment expressing prescriptions that has binding force on the agent herself.

What explains this connection between the personal imperative and motivation is that the moral agent who feels obliged to follow through on her moral judgment and actually donates to charity, seems to have this feeling of obligation (which, perhaps controversially, I will simply understand it as a form of motivation) *because she has engaged in some sense with the moral considerations at stake*. This can be understood as a difference in commitment level by the agent towards the moral considerations at stake. Take for example the following consideration. There seems to be this dichotomy between an agent-centered conception of practical reason and an other-regarding conception of morality, as noted by Paul Hurley:

[T]he demands of practical reason are determined from the personal standpoint, the standards of morality from the impersonal standpoint... to demonstrate the rational

---

<sup>75</sup> Svavarsdóttir 2006: 171. One must be careful not to interpret ‘follow direction’ here to imply that accepting something as one’s own guide leads to actual action. Svavarsdóttir only has in mind that acceptance as one’s own guide is connected to being moved to follow its directions.

authority of moral standards... is to show that there is the requisite sort of linkage between the impersonal and the personal standpoints such that the properly trained agent has decisive reason, from the personal standpoint... to conform his actions to moral standards of rightness determined from the impersonal standpoint.<sup>76</sup>

If we grant that this is the correct conception of both practical reason and morality (which of course does not have to be granted but are both fairly common-sense conceptions of those terms), my proposed account presents a solution to how practical reasoning and our understanding of morality is connected to explain the relationship between moral judgment and motivation. That is, if an agent, through practical deliberation, is motivated from her moral judgment, it makes sense to say that the agent must have related a moral consideration that should be perceived as agent-neutral, into one that is essentially agent-relative, namely relative to herself. Thus what I propose is that this personal imperative should be understood as an essentially agent-related version of a moral imperative.

In summarizing my proposal of how to understand the source of moral motivation then, I argue that even though our practical rationality can sufficiently agree that something is a worthwhile moral guideline, this does not mean that practical rationality will necessarily always align with morality insofar as motivation is concerned. One reason offered was that general moral judgments of the form 'Φ is right' seem reasonably to be first and foremost a judgment about what is the correct thing to do, and thus likely to have evaluative authority in the absence of practically prescriptive authority, and hence of motivation. It sounds plausible to me to understand the relationship between motivation and the type of judgments most aptly understood as personal imperatives as a necessary connection. However, as I have argued so far throughout this chapter, a personal imperative might not necessarily follow a 'one ought to Φ' moral judgment (a judgment

---

<sup>76</sup> Hurley 2006: 688.

implied by a ‘ $\Phi$  is right’ moral judgment in a rational agent), and so motivation seems to have a contingent connection to moral judgment generally speaking. What is thus required for motivation to occur is that a practically rational agent needs not only to agree with the authority of the moral considerations, but to align *herself* with it as saying that those considerations are something that *she* needs to follow, that there is a reason for *her* to follow it, with the distinction between moral judgments that do and do not motivate being understood as a difference in commitment level by the agent.<sup>77</sup>

#### **IV. Can Amoralists Really Make Genuine Moral Judgments?**

One of the aspects of my proposed view is that an agent can make a moral judgment without being motivated (hence endorsing externalism). Any internalist would reject this proposal right away and, as I briefly mentioned in 1.4, would simply explain away this aspect of my view as that the amoralist is not making a sincere moral judgment but instead is making a judgment in the inverted commas sense. Since the possibility of holding two moral beliefs – one that can motivate and one that does not – is a key aspect of my view, in order to strengthen my proposal I address this objection in the remainder of my paper before I finish with my closing remarks.

In discussing the two ways an agent can hold a moral belief, I have been using the moral expressions ‘ $\Phi$  is right’ and ‘I/one ought to  $\Phi$ ’. It actually seems plausible to me that these expressions can be made by agents from multiple different perspectives and with various phenomenological experiences. Some may fall under what the internalist would consider a sincere

---

<sup>77</sup> I admit that in my account I have failed to mention any sort of demarcation line for how deep of a level of commitment is required by the agent for motivation to occur. Frankly speaking, I do not have an answer other than to say that I think that this might be a substantive matter requiring empirical analysis. Therefore I am open to and welcome criticism regarding this point.

moral judgment (entailing the presence of motivation as a prerequisite) while some would be understood as insincere (or inverted commas) judgments regardless of the fact that the mental states contain normative content. But because I believe that an agent's perspective and phenomenological experiences from whence they utter such moral statements can vary so greatly, I think the dynamic is not strictly limited to a 'black and white' case of either being or not being a real or insincere moral judgment. Rather, what I personally think is that moral expressions are highly complex statements and there could be a number of various psychological states that can in some way be related to the normative content of such expressions. However, this is not the place for such a discussion (and honestly speaking, I do not have enough background in this area to carry such a debate forward). Thus I restrict myself to strictly dealing with the inverted commas objection as it used by internalists against the plausibility of making a genuine moral judgment without being motivated. I begin with an introduction and initial response to the internalist's justification for claiming that amoralists cannot but be making moral judgments in an inverted commas sense.

#### 4.1. The Inverted Commas Objection

Sometimes the inverted commas objection seems to amount to the idea that an amoralist is (i) trying to deceive the listener by just coming off as making a genuine moral judgment without sincerely believing in what she is claiming. More commonly though, as partially alluded to in footnote 19, the idea is that (ii) the amoralist is somehow ignorant of the fact that she is not really making a genuine moral judgment – and that she is rather judging from the viewpoint of what she believes other people, or society, would judge in regard to said moral content. Of course such acts are possible: an agent could, for whatever reason, try to deceive the listener by making an insincere moral judgment; as well as in other occasions, such an agent could really just be ignorant in the sense that she does not have the capacity to truly understand the concepts 'right' and 'wrong' which

she nonetheless continues to use. Claim (ii) is the more serious contention against the externalist. In the following sections, I deal with each claim separately. First, I show that claim (i) is not a serious threat to the externalist assertion for the plausibility of the amoralist. I then discuss Smith's variation of claim (ii), as I believe it is a stronger version of this particular contention, and provide my arguments against Smith's criticisms.

One thing to note is that it seems that in both cases, the agents can be correctly understood as expressing moral statements. After all, the statements contain moral phrases which are uttered by the agents. However, there seems to be this commonsense assumption that moral motivation normally, perhaps somewhat reliably, follows sincere moral judgments. This assumption has strong phenomenological support, and arguments such as Smith's depiction of the 'good and strong-willed person'<sup>78</sup> seem to also reflect that this is an assumption that cannot be easily dismissed. Against this backdrop then, it seems that the putative amoralist is a case of deviance that needs to be explained away because it does not match with our ordinary moral experiences and their parallel psychological explanations. Thus the resulting consequence is explaining them away as not being moral judgments proper. The problem though is that *proper* in this instance is then understood as reflecting the 'usual' intuitions of what constitutes moral judgments (reflecting the assumption).

Relatedly, the conflict between internalists and externalists arises when the internalist claims that the inverted commas use is universally the case for every instance of a putative amoralist. If that is the claim, then the internalist objection seems quite implausible to externalists in light of the fact that it would require for the internalist to have a precise understanding of the psychology of every amoralist. But the fact of the matter is that the internalist does not make this claim coming from some insight regarding the psychology of an amoralist, but rather from what they accept as a conceptual truth in regard to the connection between moral judgment and

---

<sup>78</sup> Smith 1994: 71-75.

motivation, with such a view disregarding the plausibility of an amoralist right from the start. Thus, even if the conception of an amoralist is patently false, the internalist argument fails to show that this is the case unless they can somehow provide an adequate psychological account of the amoralist which reflects their view. Thus it seems to me that the internalist claim of the necessary connection between moral judgment and motivation *as a conceptual truth* rests on substantial *a posteriori* assumptions about our actual moral experiences. Of course there is nothing wrong with making and trying to understand our *a posteriori* assumptions about our moral experiences. However if this is the case, then the internalist claim regarding the relationship between moral judgment and motivation cannot be served as a substantial *a priori* constraint on theories of moral judgment.

#### 4.2. Insincere Moral Statements

To say that an amoralist's moral judgment is simply a form of trying to deceive the listener, does not really appear to be what either the internalist or externalist are wanting to refer to insofar as what seems to be entailed by making a moral judgment. After all, if the purpose of the amoralist expressing the moral statement is to have someone else follow it without the agent having any inclination to sincerely follow it herself, then this seems to simply be an insincere moral statement; one that might not even require for the amoralist to have a genuine conceptual grasp of the moral terms involved.

Suppose Lucy and her best friend, Mina, are attending a music festival together. Lucy knows that Mina has a drinking problem and is in a recovery program. Lucy also enjoys drinking very much and drinks on a regular basis but not to a degree that would be considered irresponsible. During the festival both Lucy and Mina are offered complimentary cocktails by a waitress. Lucy, sensing that Mina is visibly tempted, says "it is (morally) better not to drink alcohol" without really

thinking there would be a problem with having just one glass. Here, Lucy is most likely expressing an opinion which she does not personally agree with. If so, we can consider this a case of a person truly making an *insincere statement*. The specific fact that Lucy does not want to put Mina in the agonizing position of having to watch her drink while standing right next to each other, causes Lucy to utter a moral statement which she does not really endorse.

We can see here that, while there is an obvious connection and consistency between Lucy's moral statement and her behavior (she continues to refrain from drinking in front of Mina), this should not be construed as reflecting the internalist position. For the internalist contention is that the moral expression must reflect a genuine belief in the judgment (which under my proposal, the epistemic commitment would suffice to satisfy this condition). But that is clearly not the case in Lucy's situation. After all, Lucy would most definitely have no problem drinking when she is not around Mina's presence and does not have to worry about her feelings. Since Lucy is not actually giving her approval to the content "it is better to not drink alcohol", there is no normative content that would move her to act accordingly were she in 'normal circumstances'.

In other words, we can see that for Lucy there was a clear reason to utter a moral statement which she does not endorse. This led to her 'deceiving' both the cocktail waitress as well as her friend Mina. If we were to ask her how she felt about drinking alcohol in a responsible manner, she would most likely express a statement that is completely different from what she uttered in front of Mina, as long as she is not in a similar situation. So the internalist has every reason here to assume that Lucy is making an insincere moral judgment, but the externalist would have no problem also accepting this claim as Lucy's purported moral expressions would not count as what the externalist would consider a real moral judgment either, let alone a sincere one.

However, what if Lucy actually *insists* that drinking alcohol is bad but she just simply does not care and is not motivated to act correspondingly? What if the situation was changed to where



Lucy never had a friend like Mina and where Lucy is more of a moderate drinker who might drink occasionally or at social gatherings but even though she responds the same – that drinking alcohol is bad – she nevertheless does not hesitate to drink when the occasion arises? We no longer have the reason that we had to allow us to assume that she was making an insincere moral judgment and the only factor that seems to indicate that she might not be agreeing with her moral statement (besides the fact that she does not completely abstain from drinking) is simply that she *lacks the motivation*. The internalist answer would presumably be the same, that Lucy cannot be making a genuine moral judgment since she is still not motivated accordingly. However the situation has completely changed and so it would seem that, at the very least, the internalist answer fails to accommodate this particular aspect that is different from the first scenario. That is, this change in particulars calls for a reassessment of whether Lucy is, as a matter of fact, still making the same *inverted commas* response or not; but for an internalist, reassessment is unnecessary since their answer will simply correlate accordingly to whether she is motivated or not. How would a third-party observer conclude of Lucy's moral judgments?

#### 4.3. Non-Cognitivist Versus Cognitivist's Use of the Inverted Commas Objection

One way an observer might conclude is in a form of the second internalist response, that Lucy does not really understand the concepts involved in the normative content and so does not endorse the moral content that it is better to not drink alcohol when she utters those words. Rather, what she is doing is likened to merely remarking upon a sociological observation than a sincere statement of one's convictions; much like how an anthropologist would report on the idiosyncratic moral sentiments of an indigenous tribe located remotely in the jungle. This is the contention made famous by R.M. Hare.

Hare is known as a prescriptivist, a form of non-cognitivism. Non-cognitivism is the ideal position from which to argue that amoralists are uttering moral expressions from an inverted commas sense – since under their view, moral judgments are identical to motivating states and so a necessity of expressing a sincere moral judgment is to be motivated. I will not argue against the non-cognitivist's charge against the amoralist making an inverted commas judgment because I have taken a cognitivist position regarding what moral judgments are (and to be honest, I am not even sure I can provide an adequate response other than argue that moral judgments should be construed as cognitive mental states rather than simply expressions of approval or denial). So in the remaining sections, I discuss a cognitivist internalist's charge against the amoralist as making an inverted commas judgment, namely Smith's charge. Before going into Smith's criticism specifically, let me provide some preliminary thoughts regarding the general cognitivist internalist charge of the amoralist's inability to make sincere moral judgments.

Unlike the non-cognitivist, the inverted commas argument seems to lose some of its force when considered from the cognitivist position because of what this position entails regarding the make-up of moral judgments. If, under the cognitivist view, moral judgments purport to be mental states which first and foremost provide a description of certain normative features of the world which then somehow have the ability to motivate agents that are in cognitive contact with said contents, then (reflecting Smith's internalist view from chapter II) the most likely explanation for the manufacturing of motivation comes from a second aspect of the cognitivist position; that moral beliefs are a type of evaluative beliefs where the agent places *value* on the normative content of the moral judgment. And it is from this valuing of the normative content – or the acknowledgment of the value of the normative content tied in to a specific internalist conception of practical rationality – that motivation occurs.

Now if valuing were simply a form of desiring then the cognitivist picture would be fairly easily explainable as well. The moral judgment generates the necessary motivating state of valuing required for motivation. But, as was already analyzed by Smith himself, valuing cannot be reduced to a form of desiring. If valuing was conceptually reducible to some form of desiring, then it would seem impossible for an agent to take a critical stance against that which she desires. However, there are plenty of examples in the literature (as well as in real life) that serve to deny such an analysis. Consider the recovering drug addict who has a desire for another dose of the illegal substance yet does not value doing so as she is sincerely trying to clean up her life. Inversely, there is the example of a struggling but recovering drug addict who tries hard to align her desires to her newly rediscovered values of living a clean and sober lifestyle.

One might offer the explanation here that what the struggling drug addict is doing is simply trying to line up her first-order desires to her second-order desires. But this would not serve the internalist as Nagel (himself a cognitivist internalist) has brought up an important distinction between such differing types of desires and that the latter type is best construed as a mental state that possesses cognitive status since they have been derived from our beliefs about what matters (or what is valuable).<sup>79</sup> Thus, it seems more appropriate to analyze valuing as a cognitive mental state, which can be aptly explained as a mental state involving cognitive reasoning processes to assess the worth of something through means of considering its strengths and weaknesses, adducing various reasons and judging their relations to one another, bearing in mind less salient aspects, etc.<sup>80</sup>

So the cognitivist internalist's use of the inverted commas argument has a burden that the non-cognitivist does not have to bear. We can see this burden from another perspective as well. As already indicated above, cognitivists identify moral judgments as essentially pertaining descriptive,

---

<sup>79</sup> I do not have the space here to go into the specifics of this well-known distinction between what Nagel calls 'unmotivated' and 'motivated' desires. See Nagel 1970: 29.

<sup>80</sup> Shafer-Landau 2003: 126.

semantically meaningful and epistemically assessable content. In other words, moral judgments are such that they concern certain matters of the world which remain the same no matter the accompanying attitudes of agents who make them. As such, the assertion that two people in the exact same circumstances, whose moral expressions have exactly the same content, are *meaning* two different things, strictly because one is motivated and the other is not, does not seem very convincing (hopefully it is obvious that my cognitivist *externalist* view would be a solution to this problem as distinguishing between two ways of holding a moral belief can serve as an explanation to this scenario). Nonetheless, Smith provides a compelling inverted commas objection coming from the cognitivist internalist perspective.

#### 4.4. Smith's Inverted Commas Objection

Noting an objection from Brink, Smith has “misgivings, at least as regards the details of Hare’s version.”<sup>81</sup> He believes that the sentiment behind Hare’s inverted commas objection is on the right track but needs to be refined in order to really be able to defend against the externalist assertion for the plausibility of the amoralist. Smith offers the following explanation as to what the inverted commas objection should consist of. Roughly speaking, Smith believes that it is not ad hoc for internalists – or in his case, ‘defenders of the practicality requirement’ – to claim that amoralists cannot be making genuine moral judgments even if this is, *at most*, all that the internalist can claim. Smith’s inverted commas objection amounts to the assertion that a person who makes a moral judgment without being motivated lacks sufficient ‘mastery of moral concepts’ and is therefore not making a genuine moral judgment.<sup>82</sup> That is, an amoralist cannot help but fail in making a sincere moral judgment because such an agent, who uses moral terms and lacks motivation, does not really possess a genuine understanding of the concepts belonging to the terms involved. In order to see

---

<sup>81</sup> Smith 1994: 68.

<sup>82</sup> Ibid. 68-71

why this is not an ad hoc claim, Smith explains what holding mastery of moral concepts amounts to via an analogy regarding what is entailed by having mastery of color terms.

Suppose a person blind from birth nonetheless has a completely reliable facility to use color terms accurately. Would we ascribe to this person genuine possession of color concepts? The argument progresses towards what the conditions are for being able to attribute genuine possession of color concepts. For example, both the internalist and externalist would accept reliable usage as being one necessary condition. Where the disagreement lies is that the internalist also maintains that visual experience is a prerequisite, whereas the externalist denies this as being necessary. Of course visual experience here is understood as being analogous to moral motivation. While the color analogy, on its own, is simply an appeal to one's intuition<sup>83</sup>, the point for Smith is that, just as visual experience seems to be a necessary condition in order to attribute to someone mastery of color concepts, moral motivation seems to be a necessary condition in order to attribute to someone mastery of moral concepts – that is, for an agent to be able to make genuine moral judgments.

However, to say that this is Smith's main point regarding the color analogy would be unfair to Smith's contention, as he acknowledges that the color analogy by itself cannot seal the victory for internalism. Rather, the conclusion Smith wants to draw is that the externalist cannot appeal to the blind person seemingly being able to use color terms reliably as evidence that amoralists are plausible; for this would simply be begging the question against the internalist assertion of the necessary connection between moral judgment and motivation. In other words, simply denying that visual experience is a necessary condition for a blind person (or anyone for that matter) to be able to make real color judgments – that is, have genuine possession of color concepts – is simply presupposing externalism, and not arguing for it. One thing to point out here is that at the same time, the externalist could likewise claim that the internalist assertion, that motivation (visual

---

<sup>83</sup> For Smith's internalist argument, see Smith 1994: 68-76. For an externalist response, see Brink 2008: 294-298.

experience) is a necessary condition for possession of moral (color) concepts, is begging the question against the externalist's contention of the plausibility of the amoralist. Either way, I believe we have grounds to question Smith's version of the inverted commas objection, primarily his assertion that an amoralist lacks sufficient 'mastery of moral concepts'.

First off, lacking sufficient mastery of moral concepts could be misinterpreted as in a sense being a charge of *unintelligibility* on the amoralist, at least insofar as moral terms are concerned. Consider first that Smith differentiates his inverted commas objection from that of Hare's 'sociological observation' version because he notes that:

[T]here seems to be nothing incoherent about the idea of an amoralist who claims to have special insight into what is *really* right and wrong; an amoralist whose judgments about what it is right and wrong to do are therefore, even by her own lights, out of line with the judgments of others. But if this is right, then the judgments of amoralist can hardly be thought of as judgments about what other people judge to be right and wrong.<sup>84</sup>

So then, if an amoralist utters terms such as 'right' and 'wrong', and she seems to be using these terms in appropriate context in whatever circumstances in which she finds herself using them in – and the amoralist is also not to be understood as speaking of 'right' and 'wrong' in the sociological observation sense – then what is it that the amoralist agent is supposed to be meaning here precisely? It seems that such an agent cannot be meaning anything at all – that she is uttering something akin to nonsense. But if this is the sense in which Smith's contention as a charge of unintelligibility is supposed to be understood, it seems much less convincing to argue that an amoralist just utters nonsense, rather than that the amoralist is making a real moral judgment but just happens to not be motivated for some specific practical reason. After all, concepts such as 'right' and 'wrong', while

---

<sup>84</sup> Smith 1994: 68.

abstract, are not particularly difficult concepts to understand in an everyday practical usage sense since we are all so familiar with these terms already.

Furthermore, I question whether his color analogy even works at all. The reason I think that it might not go through is because visual experience and motivation are quite different mental states. Whereas even Smith would agree that motivation is a non-cognitive state, visual experience is most commonly understood as a cognitive state, one that purports to represent a certain aspect of the world.<sup>85</sup> Relating the color example to what it means to have mastery of moral concepts would then be asking the externalist to accept an analogy where the comparable mental states are not categorically the same. Therefore, it seems conceivable that the externalist can grant Smith that visual experience is a requisite for having mastery of color terms, while maintaining at the same time that motivation is not required in order to have mastery of moral terms. However, for the sake of argument, I bracket this potential inconsistency aside.

#### 4.5. A Rebuttal to Smith's View

The idea that an amoralist *lacks* sufficient mastery of moral concepts seems to imply that it is due to some *incapacity* that she is unable to understand moral terms appropriately enough to make sincere moral judgments. But recall at the end of section 1.4 I stated that an amoralist need not be someone who is never motivated by a moral judgment. To fail to be motivated on just one occasion suffices to be properly termed an amoralist. What this also implies is that such a person would not be labeled an amoralist *until* that first point in time when she makes a moral judgment without being motivated. If so, then consider a 'recently turned amoralist': someone who has always been motivated by her moral judgments just like any other moral agent, until one day she has what she perceives to be (in the loosest sense) an epiphany of sorts regarding the authority of

---

<sup>85</sup> However, even this point needs qualification as some might suggest that it is, at best, controversial whether visual experience is representational.

morality. It is not that her moral compass has changed, she continues to make the same discernments and draw the same conclusions using the same moral terms as before, except that it is no longer the case that these judgments are accompanied by motivation.

How would Smith respond in regard to such an agent under the definition of the amoralist he has provided? Would he say that such a person either (i) never had mastery of moral concepts, or (ii) lost it at the moment she stopped being motivated? It would seem that neither answer is very appealing. If Smith chooses (i), then he would need to provide a plausible account of how someone who does not have a genuine understanding (mastery) of moral terms can nonetheless be motivated accordingly; or Smith is essentially admitting that one could make genuine moral judgments at some point in one's life without ever having mastery of moral concepts. However, if Smith concludes (ii), that this agent somehow lost mastery of moral concepts the moment she stopped being motivated, that would give rise to a whole set of issues on its own. First off, such a claim would bring about questions regarding whether someone could lose mastery of a concept after the fact that the agent has already acquired it (excluding instances of mental diseases, such as Alzheimer's, playing a primary cause); or what exactly *mastery* entails (as mentioned above, I would be inclined to think that Smith would not want to commit himself to the position that if one does not have mastery of moral concepts, then the moral expressions of such a supposed amoralist is akin to uttering nonsense). Second, and more crucially, whatever answer Smith provides would still be unsatisfactory in light of the fact that whatever it was that caused her to no longer be motivated, at the very least we can still ascribe to her continuity and consistency in her cognitive faculties.

Relating this to Smith's color example, even if we grant that visual experience is a necessary condition of mastery of color terms so that a person blind from birth does not have mastery of color terms, if a person went blind in their adult life, and if that person is able to reliably



use color terms even after she went blind, I would imagine that we would have to attribute to her mastery of color terms, seeing as how she did have plenty of visual experiences of what blue, or red, or yellow is like prior to going blind. The fact of the matter is that this blind person can not only use color terms reliably, but when she uses color terms in her speech – such as uttering the phrase ‘please grab the red jacket for me’ – she is not intending here to mean ‘please grab the jacket which society or other people have deemed to be red’, but actually intending to mean ‘please grab the jacket which as a matter of fact is red’.<sup>86</sup>

Finally, from an observational standpoint it seems that Smith’s inverted commas objection is counterintuitive. Imagine someone who became colorblind at a certain point in time during his adult life but nobody knows about it. She continues to point out the same color, red, in her expressions depicting apples, street signs and other objects which are as a matter of fact red as if she were no different from any non-colorblind person. It would seem that, simply from observation without being privy to the fact that she is colorblind, Smith would be inclined to judge this agent as having possession (mastery) of color concepts. It is only when she reveals that she is colorblind, and therefore does not have the requisite sense data to confirm whether the objects she is referring to actually possess the color red or not, that Smith would change his stance regarding whether the agent has possession (mastery) of the color concept red.

Of course it can hardly be said that this (hypothetically) mistaken identification by Smith would therefore overthrow his claim. Smith argues for a conceptual connection and therefore it would not suffice for a mistake from an observational standpoint, or any empirical evidence to the contrary, to undermine Smith’s argument even if it was made by Smith himself. What this argument does, however, is put the internalist at a disadvantage from the empirical perspective. Remember, the internalist argues for the *conceptual impossibility* of an amoralist, whereas the

---

<sup>86</sup> Brink (2008: 296-297) makes a similar point.

externalist simply needs to show one instance where such a conception is possible. Therefore, even if a thousand instances reflecting internalism are witnessed; a singular, *genuine* instance reflecting externalism (the amoralist) would suffice to push any empirical argument the externalist way. So when considering which side, the internalist or externalist, seems to be begging the question on the other, it seems to me that there is more reason to think that the internalist is begging the question against the externalist. If nothing else, at the very least I think we can conclude that the burden of proof is greater on the internalist.<sup>87</sup>

#### 4.6. An Amoralist's Judgment is the Same in *Kind* as Those of a Moral Agent's

In differentiating his inverted commas argument from Hare's, Smith mentions that "the point is not that amoralists really make judgments of some other kind: about what other people judge to be right and wrong, for example. The point is rather that the *very best* we can say about amoralists is that they try to make moral judgments but fail."<sup>88</sup> The phrase 'not...judgments of some other kind' is what interests me. Let me try and expand upon this point as far as what Smith might have in mind here. Possibly Smith could make the following statement, 'It might be difficult to provide an account of how exactly we should understand the amoralist's psychological state. But we can be certain of two things: one, it is not a genuine moral judgment; and two, it is not a psychological state that differs in *kind* from that of a genuine moral judgment.' Suppose that this is Smith's contention and that he is correct: an amoralist's judgment is not the same as a moral agent's judgment, but it does not differ in type from such moral agent's judgments. Not only do I believe that this claim is plausible, but I also believe that this is precisely the kind of picture that accommodates my positive account on the relationship between moral judgment and motivation.

---

<sup>87</sup> See Svavarsdóttir (2008: 316-320) for a similar objection to Smith.

<sup>88</sup> Ibid.

So bearing in mind that amoralists do not make the same moral judgments as moral agents, but they are the same in ‘kind’, an amoralist’s judgment must miss some aspect or element that differentiates her judgment from that of a moral agent’s. Presumably, the amoralist’s psychological state in which she is in contact with the normative reasons that are candidates as moral facts lacks something that the state in which the moral agent is in contact with the same normative reasons has. The question is are these elements essential to what should be considered a genuine moral judgment, or are they additional features? If we grant that this question cannot be easily answered due to the differing intuitions of internalists and externalists, then I propose my view as a plausible way to answer this question.

Following my proposed account, it would be the case that the amoralist’s judgment is missing the element of the agent being *personally engaged* with the moral content and so not seeing the moral imperative as a personal imperative. I believe that this aspect should not be considered as an essential element of what constitutes making a real moral judgment. Yes, the normative content of moral judgments applies to each and every individual, and thus each and every individual should take a personal interest in the moral judgments they make. But at the same time, if we are to uphold the objective and universal nature of the normative content of the moral judgments – that they apply to everyone equally, not just especially to Tommy or Sally or Lucy or whoever happens to be engaged with the contents at that time – then I do not think that the personalization aspect should be considered an essential element of what constitutes a genuine moral judgment. This should be particularly true for moral realists, who would assert that moral facts exist independently of rational agents (I do not want to imply that I commit myself to moral realism, I am just making a point that seems important).

Let me end by posing one last question regarding Smith’s general cognitivist internalist view as a final criticism. If Smith’s internalist assertion of the *necessary* connection between moral

judgment and motivation is to be understood as a conceptual claim<sup>89</sup>, and if Smith's cognitivist understanding of the etiology of motivation is that the content of a moral judgment *causes* (or generates) the motivation<sup>90</sup>, then how should we make sense of how the connection between moral judgment and motivation is *both* conceptually necessary *and* causal at the same time? For it seems to me that if the connection was conceptually necessary, it means that one cannot exist without the other. But then how could a moral judgment exist prior to it generating motivation? Further, should a moral judgment at its conception (just prior to it producing the motivation) be considered an insincere moral judgment since it has yet to be connected to motivation at that stage? Perhaps this question erroneously makes use of the conception of time-ordering and perhaps there is something regarding the notion of causality that I have yet to understand (I admit that I am not nearly learned enough in either of these concepts from a philosophical perspective). Or perhaps this is a valid question that a cognitivist internalist view such as Smith's in particular has to deal with. Either way, I welcome any kind of answer by a cognitive internalist and such answers could very well help as a response to my criticisms against Smith's inverted commas objection as well.

### **Closing Remarks**

The goal of my master's thesis was to provide a comprehensive, perhaps somewhat unconventional yet hopefully convincing, solution to understanding the relationship between moral judgment and motivation. The position I forwarded took externalism and non-Humeanism as its framework and argued for an intrinsic yet contingent connection between moral judgment and motivation. In order to support this externalist view, I first started off in chapter II by targeting an internalist account which I felt was most adequately contrasting to mine; that being Smith's cognitivist internalist view. I laid out what his view amounted to and provided my arguments

---

<sup>89</sup> Smith 1994: 61.

<sup>90</sup> Ibid: 73.

against this form of internalism throughout chapter II. Initially I argued that Smith's account failed to make a certain type of distinction within practical judgments in the form of *seeing* the best means to one's ends versus *pursuing* (or intending) the best means to one's ends. I argued that these are two importantly distinct forms of mental acts which do not necessarily always coincide. Furthermore, this lent support to the possibility that we can have doubts about the rational authority of morality and that such doubt can lead to failure in motivation. In the latter half of chapter II I forwarded what I believed to be the key problem with internalist accounts such as Smith's: that they conflate the action-guiding nature of morality – that is, what it means for something to be normative – with moral motivation.

Chapter III consisted of the layout of my positive account. As a starting point to my positive account, I proposed that there is more than one way an agent can hold a moral belief. I distinguished between moral judgments of the form 'Φ is right' and 'I/one ought to Φ'. I supported my distinction by integrating a moral variant of Moore's paradox as an analytical tool. The key idea of using this moral variant was to show the possibility of two ways of holding a moral belief that are distinct and so could reflect the mindset of an amoralist. I did so by positing the terms 'acknowledging' and 'endorsing' respectively to each half of the conjunct of a moral version of a Moorean paradox to serve as the key references to the 'belief' terms so that the statement is no longer a paradox, but a solution of how an unmotivated person could hold such opposing views. Further, focusing on the latter (motivationally conducive) form of 'I/one ought to Φ', I tried to provide an explanation of how motivation occurs from the utterance, or holding, of such a moral belief by an agent. I argued that the operative state was turning an impersonal moral imperative into an agent-relative personal imperative. That is, motivation occurs when an agent *personally engages* with the moral considerations at stake.

In the first half of chapter IV, I introduced the inverted commas objection levied by internalists upon the externalists positing of an amoralist. The discussion revolved around whether it was possible for an amoralist agent to genuinely make a real moral judgment; and if not, then what was meant by their so called ‘moral judgments’. Here I argued that the internalist use of the inverted commas objection was based on what they conceived of to be a conceptual truth regarding the relationship between moral judgment and motivation, when what they needed was to provide a psychological account for why it is the case that the amoralist cannot help but utter moral judgments in the inverted commas sense. I further argued that even their contention of the necessary connection between moral judgment and motivation as a conceptual truth was not an *a priori* argument at all, but based upon *a posteriori* assumptions regarding our moral experiences. After my criticism of the inverted commas objection, I provided an example of what should be considered an insincere moral judgment and how it differs from what the externalist is asserting regarding the amoralist agent who is unmotivated by her moral judgment. I then further reduced the domain of the objection to just the cognitivist argument of the amoralist making moral judgments from an inverted commas sense as a transition to my final chapter.

The latter half of chapter IV was dedicated towards describing and arguing against Smith’s cognitivist internalist view and his use of the inverted commas objection. Smith’s inverted commas objection amounted to the idea that amoralists fail to make genuine moral judgments because they lack ‘sufficient mastery of moral concepts’. In support of his claim, Smith uses an analogy regarding the use of color terms to argue that, just as visual experience seems a necessary condition for mastery of color terms, so does motivation seem to be a necessary condition for mastery of moral terms. I responded that even if we grant that visual experience was a necessary condition for mastery of color terms, his analogy still fails to account for the amoralist agent who once used to have the necessary condition satisfied, but through some event, no longer does while maintaining all

other cognitive abilities the same. I further tried to analyze what Smith's lack of sufficient mastery of moral terms could be meant to infer and one of the possible conclusions was that the amoralist agent is expressing utter nonsense. I surmised that if this was the conclusion, the externalist explanation was far more convincing than Smith's as far as what the amoralist agent could be meaning by his use of moral terms.

Lastly, I pointed out one final aspect of Smith's version of the inverted commas objection – that an amoralist's judgment is not different in *kind* from that of a moral agent's – and suggested that we could expand upon this point to support my positive account of an agent holding moral beliefs in more than one way. If an amoralist's judgment is the same in kind of that of a moral agent's then what is missing is an additional feature or element which is responsible for motivation. I argued that my position could provide an answer as to what that additional element was, but from an externalist aspect where that element was not crucial to defining what a real moral judgment is. I wrap up my master's thesis with these final thoughts.

I submit that there are two primary purposes for an agent making a moral judgment. One is to declare what is the morally right or wrong thing to do. The other is to hope to influence another person's behavior by offering strong moral reasons. Insofar as the latter is concerned, we hope that people agree with the normative content of our moral reasons and that they will therefore behave accordingly. But there is a very important implication regarding why we hope that they will behave accordingly insofar as entailing whether one is truly motivated as a result of making a genuine moral judgment or not. And that is, that we hope that people will behave accordingly not simply because *we*, the presenter of the moral judgment, believe that it is correct to follow the moral judgment; but because we hope that the listener accepts that the normative contents of the moral judgment are good reasons for *them* as well – that the normative reasons are also *their* reasons for action. Reflecting Smith's rationalist account, which is perhaps counter-serving to some of the

points I have made in this thesis, we expect, amongst other things, that people will want to adhere to moral reasons because they serve as moral reasons for them and it will not accord with their sense of rational agency to act otherwise if they have sincerely accepted the moral reasons. Thus, in my mind, what makes a genuine moral judgment a moral judgment which also motivates is because when we act upon them, the moral judgment has engaged each individual agent personally – that each agent not only accepts the moral judgment, but also endorses the content of the moral judgment as something that *he* or *she* needs to follow.

As one last (counterintuitive) point to the main arguments of my thesis, I end with the thought that the interrelations between moral judgment and motivation is not as simple and direct as we would like it to be. There is a good chance that the relationship between the two is quite complex and multifaceted and presumably my proposal today barely scratched at the surface. This might then be a substantial reason for why neither side of the debate between internalists and externalists has been able to take a dominant position over the other. What I hope is that my proposal can provide a fresh new perspective that might nudge us ever so slightly in the right direction.



## References

- Brink, David O. (1986). "Externalist Moral Realism," *The Southern Journal of Philosophy*, 24.S1: 23-41.
- (2008). "Moral Motivation" in Russ Shafer-Landau (eds.), *Metaethics: Critical Concepts in Philosophy*. New York: Routledge. 280-307.
- (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Cueno, Terence D. (1999). "An Externalist Solution to the 'Moral Problem'," *Philosophy and Phenomenological Research*, 59.2: 359-380.
- Darwall, Stephen. (1992). "Internalism and Agency," *Philosophical Perspective*, 6: 155-174.
- Darwall, Stephen, Gibbard, Allan, & Railton, Peter. (1997). *Moral Discourse and Practice: Some Philosophical Approaches*. Oxford: Oxford University Press.
- Foot, Philippa. (1972). "Morality as a System of Hypothetical Imperatives," *Philosophical Review*, 81.3: 305-316.
- (2001). *Natural Goodness*. Oxford: Oxford University Press.
- Hare, Richard M. (1952). *The Language of Morals*. Oxford: Oxford University Press.
- Hurley, Paul E. (2006). "Does Consequentialism Make Too Many Demands, or None at All?" *Ethics*, 116.4: 680-706.
- Korsgaard, Christine. (1986). "Skepticism about Practical Reason," *Journal of Philosophy*, 83: 5-25.
- Mackie, John L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin Books.
- Mason, Elinor. (2008). "VIII – An Argument Against Motivational Internalism," *Proceedings of the Aristotelian Society*, 108.1: 135-156.

- McDowell, John. (1978). "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society, Supplementary Volumes*, 52: 13-29, 31-42.
- (1979). "Virtue and Reason," *The Monist*, 62.3: 331-350.
- Nagel, Thomas. (1970). *The Possibility of Altruism*. Princeton: Princeton University Press.
- Parfit, Derek. (1997). "Reasons and Motivation," *Proceedings of the Aristotelian Society*, 71.1: 99-130.
- Pettit, Philip. (1987). "Humeans, Anti-Humeans, and Motivation," *Mind*, 96 (384): 530-533.
- Railton, Peter. (1986). "Moral Realism," *Philosophical Review*, 95.2: 163-207.
- Ross, William David. (1930). *The Right and the Good*. Oxford: Oxford University Press.
- Shafer-Landau, Russ. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Smith, Michael. (1994). *The Moral Problem*. Oxford: Basil Blackwell.
- (1989). "Dispositional Theories of Value," *Proceedings of the Aristotelian Society*, 63: 89-111.
- Svavarsdóttir, Sigrún. (2006). "How do Moral Judgments Motivate?" in James Dreier (eds.), *Contemporary Debates in Moral Theory*. Oxford: Blackwell Publishing. 163-181.
- (2008). "Moral Cognitivism and Motivation" in Russ Shafer-Landau (eds.), *Metaethics: Critical Concepts in Philosophy*. New York: Routledge. 308-356.
- Wallace, R. J. (2006). "Moral Motivation" in James Dreier (eds.), *Contemporary Debates in Moral Theory*. Oxford: Blackwell Publishing. 182-196.
- Watson, Gary. (2003). "The Work of the Will" in Sarah Stroud and Christine Tappolet (eds.), *Weakness of Will and Practical Irrationality*. Oxford: Oxford University Press. 172-200.
- Williams, Bernard. (1981). "Internal and External Reasons" in *Moral Luck*. Cambridge: Cambridge University Press. 101-113.
- Zangwill, Nick. (2008). "The Indifference Argument," *Philosophical Studies*, 138.1: 91-124.

----- (2003). "Externalist Moral Motivation," *American Philosophical Quarterly*, 40.2: 143-154.

## 국문초록

# 도덕적 동기의 원천

김대현  
철학과  
대학원  
서울대학교

이 논문의 목표는 두 가지이다. 첫째, 나는 전통적이지는 않지만 타당한 외재주의적 설명을 제공할 것이다. 둘째, 나의 외재주의적 설명을 뒷받침하기 위해서, 내재주의적 입장을 반박할 것이다. 첫 번째 목표와 관련해서, 내 입장의 큰 틀은 내재주의의 **양상(modal)** 주장을 거부하고, 인지적 내재주의의 동기에 대한 **원인(etiological)** 주장을 받아들이는 것이다. 달리 말해, 나는 도덕적 판단과 동기 간의 관계가 **내재적(intrinsic)**이면서도 **우연적(contingent)**이라고 주장한다. 즉 도덕적 판단을 이루고 있는 것이 도덕적 동기의 원인이지만, 그렇다고 해서 도덕적 판단이 항상 도덕적 동기를 유발하는 것은 아니다. 두 번째 목표와 관련해서, 나는 나의 입장에 반대되는 마이클 스미스(Michael Smith)의 인지적 내재주의적 입장을 비판할 것이다.

스미스는 완전히 합리적인 행위자는 자신이 받아들이는 규범적 이유에 따라서 욕구할 것이라고 주장한다. 따라서 내재주의는 참이고, 행위에 대한 규범적 이유를 인정하지만 적절히 동기 부여되지 않는 모든 행위자는 어떤 의미에서 실천적으로 비합리적이다. 나는 도덕적 판단이 합리적 행위의 중심 요건인 일관성 조건을 항상 충족시키는 것은 아님을 보여주는 논증을 제시해서, 도덕적 판단이 합리성의 규범들에 의해서 제약된다는 생각에 이의를 제기할 것이다. 또한 나는 스미스의 입장과 같은 내재주의적 입장이 도덕의 행위-지침적 성격과 동기를 혼동하는 오류를 저지르고 있다고 주장할 것이다. 달리 말해, 그 입장이 어떤 것이 규범적이다라는 것의 의미와 어떤 것이 동기부여적이다라는 것의 의미를 혼동하고 있다고 비판할 것이다.

나의 외재주의적 입장을 옹호하면서, 나는 도덕적 판단을 내리는 숙고 과정에는 적어도 두 가지 형태의 결정이 있다고 주장할 것이다. 하나는 (도덕적으로) 최선인 것이 무엇인지에 대한 판단으로 이해될 수 있다. 다른 하나는 (도덕적으로) 최선인 것을 실천할 것이라는 판단으로 이해될 수 있다. 나는 이 구분을 더 자세히 설명해서, 후자의 판단만이 동기를 **형성**한다는 것을 보일 것이다. 그리고 나는 이러한 도덕적 판단에 속해 있는, 동기를 형성하는 과정이 **행위자-중립적인 명령**(impersonal moral imperative)을 **행위자-중심적인 명령**(personal imperative)으로 전환시켜주는 더 강한 정도의 결심이라고 주장할 것이다.

끝으로 나는 ‘따옴표 반론(inverted commas objection)’을 다룰 것이다. 이 반론은 도덕적 판단으로부터 도덕적 동기가 반드시 따라나오지는 않는 사람, 즉 무도덕자(amoralist)가 불가능하다는 내재주의적 주장이다. 스미스 식의 따옴표 반론에 따르면, 무도덕자는 ‘도덕 개념들을 충분히 통달’하지 못한 것이다. 나는 이에 대해서 스미스의 설명이 특정한 유형의 무도덕자, 즉 어느 순간부터 무도덕자가 된 사람을 설명하지 못한다고 비판할 것이다. 또한 나는 따옴표 반론이 경험적인 사실에 주로 영향을 받아서 형성된 우리의 일부 직관과 상충한다고 주장함으로써, 나의 논증을 보완할 것이다.

주요어: 인지주의, 외재주의, 동기에 대한 흄의 이론, 내재주의, 따옴표 반론, 도덕적 판단, 동기, 비인지주의

학 번: 2012 - 23886